

学位論文 博士(工学)

インターネットにおける  
規模拡張性に優れた  
ネットワークトラヒック  
制御技術に関する研究

平成25年度

慶應義塾大学大学院理工学研究科

菊田 洸



# 目次

概要	1
1 背景	3
1.1 インターネットを支える技術	3
1.2 トラヒックエンジニアリングの必要性	6
1.3 本論文の意義	10
2 トラヒックエンジニアリング関連技術	12
2.1 トラヒック制御のためのネットワーク基盤技術	16
2.1.1 制御プレーン技術：IP ネットワーク技術	16
2.1.2 制御プレーン技術：GMPLS ネットワーク技術	23
2.1.3 管理プレーン技術：トラヒックのモデリングと トラヒックエンジニアリングの実現	31
2.1.4 管理プレーン技術：マルチメトリック環境での経路選択	34
2.1.5 伝送プレーン技術：WDM	36
2.1.6 伝送プレーン技術：イーサネット	39
2.2 トラヒックエンジニアリング関連研究	40
2.2.1 IP ベースのトラヒックエンジニアリング	42
2.2.2 GMPLS ベースのトラヒックエンジニアリング	44

---

2.2.3	下位レイヤーに関連したトラヒックエンジニアリング技術 . . .	50
2.3	本論文の位置づけ . . . . .	51
3	IP ベースのトラヒックエンジニアリングにおける 最適化計算の規模拡張性 . . . . .	66
3.1	概要 . . . . .	66
3.2	OSPF メトリック重み付け問題 . . . . .	67
3.3	OSPF 重み付け最適化の各方式 . . . . .	72
3.3.1	ILP による OSPF 重み付け最適化 . . . . .	72
3.3.2	GA による OSPF 重み付け最適化 . . . . .	75
3.4	マルチコア CPU による OSPF 重み付け最適化の並列計算 . . . . .	78
3.4.1	従来方式：単純な並列化による OSPF 重み付けの最適化計算 . . . . .	79
3.4.2	提案方式：データの依存関係を解決した効率的な並列処理 . . . . .	82
3.5	性能評価 . . . . .	84
3.5.1	並列化による計算時間の短縮 . . . . .	84
3.5.2	重み付け問題での最適化計算結果の比較 . . . . .	85
3.5.3	他のアルゴリズムとの最適化計算結果の比較 . . . . .	88
3.6	結論 . . . . .	90
4	GMPLS ベースのマルチメトリックトラヒック エンジニアリングにおける制御システムの規模拡張性 . . . . .	93
4.1	概要 . . . . .	93
4.2	マルチメトリック環境でのパス確立 . . . . .	94
4.3	従来手法 . . . . .	97
4.4	提案手法 . . . . .	101

---

4.4.1	提案手法の概要	101
4.4.2	フラッディングに基づくシグナリング手法	102
4.4.3	シグナリング転送テーブル	103
4.4.4	メッセージ追撃メカニズム	105
4.5	性能評価	106
4.5.1	フラッディングメッセージ数	108
4.5.2	ブロック率	114
4.5.3	セットアップ時間および経路選択スコア	117
4.5.4	始点ノードにおけるセットアップタイム	119
4.5.5	LSPの張り替えへの影響	120
4.6	結論	122
5	広域イーサネットにおけるトラヒックエンジニアリングのための ドメイン連続性に関する規模拡張性	126
5.1	概要	126
5.2	広域イーサネットにおけるマルチドメインVLANパス確立	127
5.3	広域イーサネットのアーキテクチャ	129
5.4	VLANパスの確立	131
5.4.1	単一ドメイン内におけるVLANパスの確立	131
5.4.2	マルチドメインにおけるVLANパス確立	133
5.5	特性評価	137
5.5.1	実験ネットワーク環境	137
5.5.2	タグスワップを実現したマルチドメイン VLANパスシグナリング	138
5.6	結論	139

---

6	広域イーサネットにおける TE のための多点間接続の規模拡張性	142
6.1	概要	142
6.2	広域イーサネットにおける多点間接続	142
6.3	P2MP VLAN パスをサポートした GMPLS 制御による 広域イーサネットのアーキテクチャ	144
6.4	P2MP VLAN パスシグナリングの実装	145
6.4.1	ピアモデルネットワークにおける P2MP VLAN パスシグナリング	145
6.4.2	オーバーレイモデルベース P2MP VLAN 経路シグナリングのための初期拡張	148
6.4.3	オーバーレイモデルにおける送信元から 始点ノードへの情報交換	150
6.5	特性評価	150
6.6	結論	153
7	総括	157
	関連著作一覧	159

# 目次

1.1	インターネットの人口の推移 . . . . .	4
1.2	インターネットのトラフィック総量の推移 . . . . .	6
1.3	最もシンプルなトラフィックエンジニアリング . . . . .	7
1.4	TEを実現する3つのプレーン . . . . .	8
1.5	各プレーンの代表的な技術 . . . . .	10
2.1	TEを実現の3つのレイヤーの各技術 . . . . .	13
2.2	IP フォワーディング . . . . .	17
2.3	OSPF によるリンクステートの交換 . . . . .	19
2.4	OSPF メッセージの例 . . . . .	20
2.5	ダイクストラ法 . . . . .	21
2.6	ECMP . . . . .	22
2.7	MPLS のラベル転送 . . . . .	24
2.8	GMPLS のレイヤー . . . . .	26
2.9	RSVP-TE によるシグナリング . . . . .	29
2.10	抽象化したネットワークの通信モデルとトラフィック行列 . . . . .	32
2.11	オフラインTEとRPC . . . . .	33
2.12	VLAN フレーム . . . . .	40
2.13	MATE: MPLS Adaptive Traffic Engineering の概念 . . . . .	46

2.14 各章の研究の位置づけと要約 . . . . .	53
3.1 メトリックと最短経路 . . . . .	67
3.2 メトリック重み付けと最短経路の変化 . . . . .	68
3.3 交叉 . . . . .	77
3.4 GA の基本的なアルゴリズム . . . . .	79
3.5 Ericsson の GA による OSPF 重み付け . . . . .	80
3.6 提案方式のデータ構造 . . . . .	82
3.7 提案方式による並列化 . . . . .	83
3.8 非並列処理と提案方式による並列処理での混雑度 L の比較 . . . . .	87
3.9 提案 GA および他のスキームを用いた混雑度 L の最適化の結果 . . . . .	89
4.1 従来の GMPLS における LSP 確立手法の概要 . . . . .	98
4.2 従来手法における LSP 確立のための 3 つのステップ . . . . .	100
4.3 提案する LSP 確立手法 (FB-RSVP-TE) の概要 . . . . .	102
4.4 シグナリング転送テーブルによるフラッディングおよび経路選択 . . . . .	103
4.5 メッセージ追撃メカニズム . . . . .	105
4.6 LSP 確立およびメトリック更新毎のフラッディングメッセージ数 . . . . .	109
4.7 LSP 要求あたりのメッセージ数 . . . . .	111
4.8 メトリック更新の度に見込まれるメッセージ発生数 . . . . .	112
4.9 提案手法および従来手法におけるブロック率 . . . . .	116
4.10 提案手法と従来手法のセットアップ時間 . . . . .	117
4.11 経路選択スコア . . . . .	118
4.12 ブロック率とセットアップタイム . . . . .	119
4.13 LSP 張り替えの例 . . . . .	120



---

5.1	VLAN フレーム . . . . .	127
5.2	GMPLS 広域イーサネットのアーキテクチャ . . . . .	129
5.3	VLAN パス . . . . .	130
5.4	単一ドメインにおける VLAN パス確立のシーケンス図 . . . . .	131
5.5	マルチドメインにおけるパス確立のシーケンスダイアグラム . . . . .	134
5.6	実験ネットワーク . . . . .	137
5.7	各ノードの実装 . . . . .	138
6.1	GMPLS 制御における 広域イーサネット . . . . .	145
6.2	P2MP シグナリング . . . . .	146
6.3	ピアモデルに基づいた標準的な P2MP VLAN パスシグナリングの例 . . . . .	147
6.4	各モデルにおける P2MP VLAN パスシグナリングの課題 . . . . .	148
6.5	オーバーレイモデルにおける P2MP VLAN シグナリング拡張の例 . . . . .	151
6.6	実験ネットワークの構造 . . . . .	152
6.7	P2MP シグナリングメッセージ . . . . .	153

## 表目次

2.1	IP ベースの TE および GMPLS ベースの TE . . . . .	41
2.2	各章における研究の課題と提案 . . . . .	54
3.1	実行速度の差 . . . . .	85
3.2	各ネットワークで 60 秒間に計算した混雑度 L の値 . . . . .	86
3.3	他のスキームとの混雑度 L の値の比較 . . . . .	88
4.1	メッセージ数を左右する主な要因 . . . . .	108
4.2	確立ブロッキングの分類 . . . . .	114

## 概要

インターネットを始めとするネットワーク技術は飛躍的に成長し、光アクセス網をはじめとするアクセスネットワーク技術の発達により、各ユーザーとネットワークは高速な通信回線で接続された。しかしコアネットワークにおけるトラヒック制御技術は不十分であり、ユーザー同士のトラヒックの輻輳により、通信効率の劣化や不規則な遅延、予測不可能なパケット損失が発生する。これらの通信品質を改善し、効率的な通信を実現するための技術として、トラヒックの経路を自由に制御するトラヒックエンジニアリングが求められている。しかしながら、インターネットのような大規模ネットワーク上でトラヒックエンジニアリングを行うには、ネットワークの規模やサービスに対する規模拡張性において課題を抱える。本論文ではこれらの課題を明確化し、それぞれに対して提案を行う。

第1章ではインターネット技術に関する特徴や問題点、そしてトラヒックエンジニアリングの必要性について述べる。第2章ではこのトラヒックエンジニアリングの実現に関係する既存の通信技術と課題を明確化し、それらの課題に対する関連研究と、第3章から第6章までに展開される各研究との位置づけについて説明を行う。第3章ではIPネットワークにおける経路最適化の規模拡張性における課題として、リンク距離の最適化計算を扱う。この最適化計算はネットワークサイズと共に莫大な計算量を要することが知られている。提案方式では並列プロセッサによりその問題を解決すべく、データ依存性に起因する並列処理のボトルネックを解

消し計算の高速化を実現する。第4章では、次世代のネットワーク制御技術である GMPLS(Generalized Multi-Protocol Label Switching) の制御システムの規模拡張性における課題として、トラヒック制御の要素となる遅延や消費電力といったメトリック情報の多様化について扱う。様々なメトリックのサポートは頻繁なメトリック更新による多量の制御メッセージを発生させ、ネットワークリソースを逼迫する。そこで、提案方式ではメトリック情報を交換する代わりに、フラッディングに基づくシグナリングにより最短経路を選択しながらパスを確立する。これによりメトリックの更新頻度に対しメッセージ数の増加を防ぐことが可能となる事を示す。第5章では下位レイヤー技術に関連した識別子空間の規模拡張性における課題を扱う。LAN 技術であるイーサネットを用いた広域ネットワークは現在注目を浴びているが、大規模ネットワークを実現するための複数ドメイン間での識別子の衝突が課題である。そこで提案方式では GMPLS により識別子を入れ換える方式を提案し、実際のプロトタイプネットワークを構築し動作を検証する。第6章ではこのイーサネットによる広域ネットワークにおいて、多点間を接続する通信の規模拡張性について扱う。提案方式では、GMPLS において多点間を接続するパスを確立するための2つの制御モデルを想定し、プロトタイプネットワークを構築して動作を検証する。最後に、第7章にて本論文の結論を述べる。

---

# Chapter 1

---

## 背景

### 1.1 インターネットを支える技術

この20年間で人類の生活を最も大きく変えた技術はインターネットである。インターネットは全世界のコンピュータをつなぐ巨大なネットワークであり、我々はインターネットを用いる事により世界中から自由に情報を集め、また世界中に自由に情報を発信することが可能となった。そしてインターネットは、今や電気、水道と同様に人々の毎日の生活を支える重要な生活基盤となっている。このインターネットの構想は1960年代からアメリカの国防総省の高等研究計画局 (Advanced Research Projects Agency : ARPA) において提唱され、開発が始まった。その後、米国内の大学や研究所に広まり、現在では世界中で何十億ものコンピュータがインターネットに接続されている [1]。そして、我が国日本においても、ここ十数年でインターネットの利用者数が急増している。図 1.1 には、日本におけるインターネット人口の推移を示す [2]。1997年のインターネット利用者は全人口の10パーセント程度だったが、現在では全人口の80パーセント近くに上る9000万人以上に至る。

このインターネットの利用形態は、インターネットの技術とともに大きく変化した。インターネットが利用され始めた当時の主な利用形態は、E-mailやWeb探索などであり、交換されるデータは主に文字データであった。そして、必要な情報を獲得する為に使われる端末はコンピュータであった。近年では動画共有や写真共有

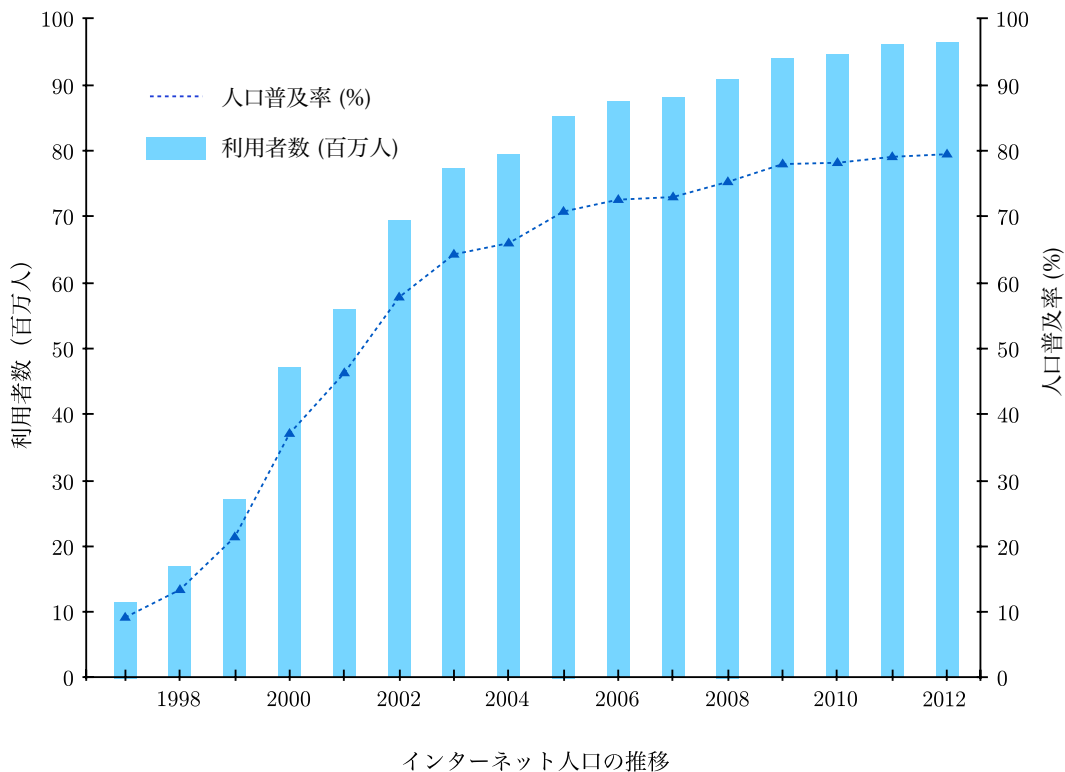


図 1.1: インターネットの人口の推移

などのサービスが一般的となり、インターネットを介して交換されるデータも大容量となった。スマートフォンなどの普及により、会話型のサービスも一般的となり、テキストチャットやVoIP(Voice over IP)による音声通信、テレビ電話などのサービスも拡大している。また、ユーザーが端末を操作する毎にサーバーとのリアルタイムな通信が必要となるオンラインゲームなどのインタラクティブなサービスも増加している。ハプティクスと呼ばれる研究分野では、ネットワークを通して触覚を伝達する手法が研究されている。

このように、インターネットは単なるデータを伝達するネットワークではなく、サービスを提供する為の通信基盤となった。これらの変化の背景に、アクセスネット

ワーク技術の発達が挙げられる。アクセスネットワークは、ユーザーとコアネットワークを接続するネットワークであり、一般的にはインターネットの通信サービスを展開するISP(Internet Service Provider)が設備する。各家庭へ光ファイバーによる高速通信を提供するFTTH(Fiber To The Home)サービスでは、2013年9月末の時点では約2500万契約に達し、100Mbpsから1Gbpsクラスのブロードバンドサービスが各家庭に提供されている。また、無線LANの発達により、ユーザーは数10から数100Mbpsの広帯域サービスを手軽に受けられるようになった。

インターネットが生活基盤となるに従い、サービスに対応した要求条件を満たす事が求められてきた。具体的には、通信品質に関係する遅延、データの損失率、ネットワークの信頼性が主なものである。特に、音声通信の場合は、音声通信パケットのデータサイズは小さなものの、遅延やパケット損失に対する要求条件は厳しい。オンラインゲームのようなインタラクティブなサービスでは、要求条件はさらに厳しくなっている。一方、動画の配信のサービスでは、データサイズは非常に大きな物となるが、多くの場合はバッファを備える事でパケット欠落時もサービスを継続させる事が可能である。すなわち、パケット損失に対するサービス毎の要求条件に差がある。しかし、現在のネットワークにおいてルーターは、通信中のデータがどのようなサービスで利用されているかを知らずに転送しているため、優先すべきデータを考慮したパケット破棄を制御することは困難である。したがって、インターネット上で通信品質を保証した通信を実現するためには、通信品質を考慮したトラヒックの制御が必要となる。

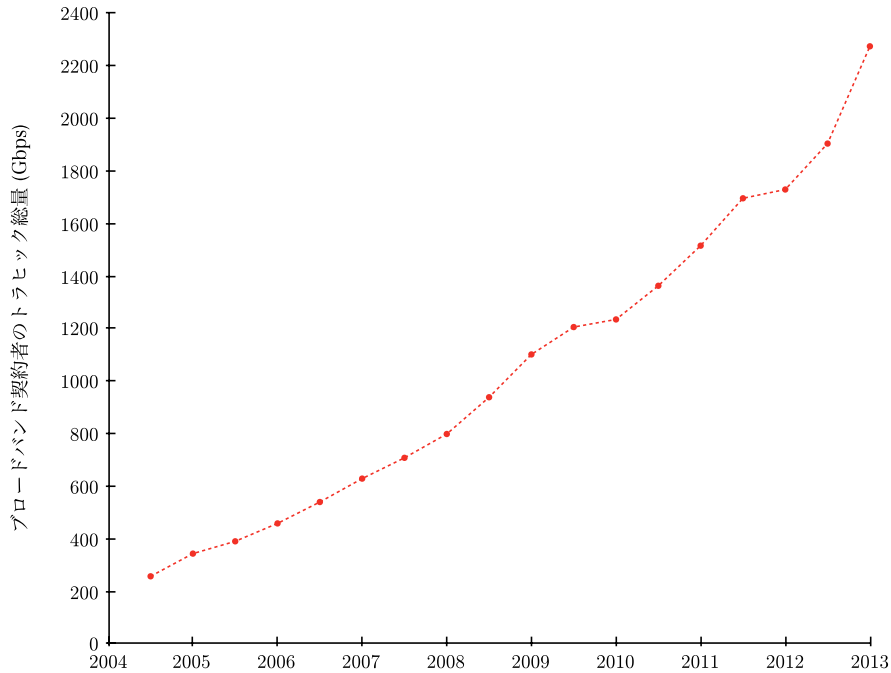


図 1.2: インターネットのトラフィック総量の推移

## 1.2 トラフィックエンジニアリングの必要性

現在のインターネットにおける通信品質の保証は、コアネットワークにおいて多くの課題を持つ。図 1.2 にインターネットのトラフィック総量を表す [3]。現在では 2300Gbps にあたるトラフィックがインターネットで交換されており、音声のような小さなサイズの packets から、動画のような大きいサイズの packets まで、様々な packets が混在している。これらの混在した packets を、今日のインターネット技術は同等に扱っている。そしてネットワーク上の混雑を考慮せずに、データの到着性だけの転送を行っているのが現実である。これらの根本には、およそ 30 年前である 1981 年に確立されて以来変わっていない TCP/IP (Transmission Control Protocol /



Internet Protocol)[4][5]に基づいた通信方式が挙げられる。この通信方式は、自律分散システムにおいてデータの転送遅延や輻輳は考慮せずに、パケットを到達させることを第一目的としていた。この方式はベストエフォートと呼ばれ、ネットワークノードであるルーターは、パケットが到着すると次のルーターへの転送を試み、転送できない場合はパケットを保持し、限界まで保持しても転送できない場合はパケットを破棄するという単純な仕組みで動作している。従って、この通信方式では、様々なトラヒックがリソースを共有するため、輻輳が発生した場合には優先するサービスのパケットの破棄を防止できない。また、トラヒックの経路を自由に制御することも不可能である。

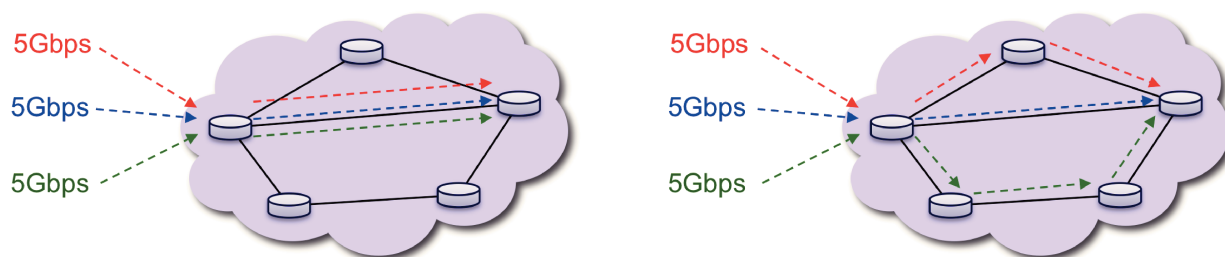


図 1.3: 最もシンプルなトラヒックエンジニアリング

これらの課題を解決する為に必要な技術が、トラヒックエンジニアリング、TE (Traffic Engineering) である。TEは多くの研究者によって研究されており [6]、本論文で扱うテーマである。TEの本質は、ネットワーク上で交換されるデータすなわちトラヒックを巧みに操作し、輻輳の回避または省電力化等、ネットワークを効率よく使えるようにする事である。もっともシンプルな輻輳回避の為にTEを図1.3に示す。図中において左のネットワークでは3つのトラヒックを同じ経路に転送しようとしている事に対し、右のネットワークではそれぞれのトラヒックを別々の経路に

転送しようとしている。この例において、どちらがネットワークを効率的に使用しているか明らかである。それぞれの通信が別の経路を通過する事により、輻輳を起こさずに通信を実現し、ネットワークを効率的に使用することが可能となる。しかし、実際のインターネット上でTEを実現する事は容易ではない。インターネットは世界中を接続したネットワークであり、様々な送信元から様々な宛先へと通信が行われている。その通信トラヒックの挙動は時々刻々と変化する。さらに、この巨大なネットワークは様々な通信プロトコル、伝送技術から成り立っており、これらの技術と整合性のあるTEが求められる。これらの既存の技術とTEを実現する為の課題や、関連する研究については、次の章において詳しく説明する。ここでは、各章の関連を示す簡単に説明を加える。

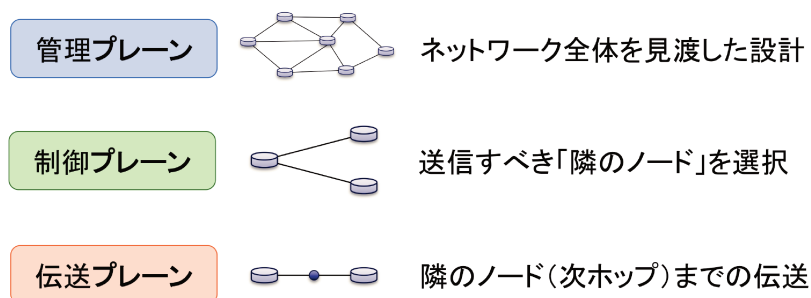


図 1.4: TE を実現する 3つのプレーン

TE を実現する技術は図 1.4 のように大きく次の 3つのプレーンに分類される。

### 伝送プレーン

このプレーンの役割はデータを隣のノードに送信することである。どの隣のノードに送るかを決定するのは、次の制御プレーンに任される。OSI(Open Systems

Interconnection) 参照モデル [7] におけるレイヤー 1 およびレイヤー 2 の主な技術はこの伝送プレーンに分類される。

### 制御プレーン

ルーティングプロトコルなどを用いて次に送るべきノードを決定することがこのプレーンの役割である。OSI 参照モデルにおけるレイヤー 3 は制御プレーンにおける重要な技術の一つである。

### 管理プレーン

TE を実現する為に、他のトラヒックの考慮などネットワーク全体を考慮したトラヒック制御の設計を行う。これは ISP のネットワーク管理者が行うか、トラヒックモニタ情報を基にコンピュータにより自律的に行う。

ここで TE の実現に対し最も重要なプレーンは、制御プレーンである。さらに、制御プレーンはパケットを転送する通信方式によって、大きく二つに分けられる。一つは、ここまでに説明したインターネットの歴史的なプロトコルである IP である。そしてもう一つは、TE を実現するために提案され、現在標準化段階にある GMPLS (Generalized Multi-Protocol Label Switching) である。GMPLS の詳しい説明は、次章で述べる。制御プレーンが IP であるか、それとも GMPLS であるかによって管理プレーンの技術は大きく変わる。これらを考慮し、本論文の各章で扱う個々の研究をマッピングした結果を図 1.5 に示す。

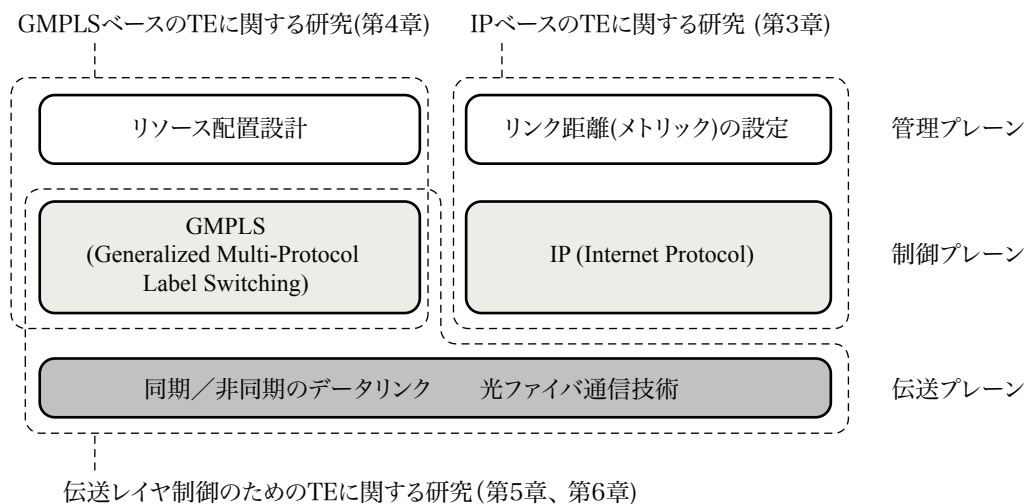


図 1.5: 各プレーンの代表的な技術

### 1.3 本論文の意義

本論文では、インターネットにおける TE 実現のための各プレーンでの課題を解決する。まず最初に第2章において、図1.5における分類に基づいたネットワークの基盤技術について説明を行い、その後、TEにおける関連研究の紹介とともに各プレーンでの課題を明確化する。そして第3章から第6章にかけて、各課題を解決するための提案を行う。そして第7章において論文を総括する。

## 参考文献

- [1] Jonathan G. Koomey, “Estimating Total Power Consumption by Servers in the U.S. and the World,” Final report, February 15, 2007.
- [2] 総務省, “インターネット利用人口の推移,” 情報通信統計データベース, June 2013.
- [3] 総務省, “我が国のインターネットにおけるトラフィック総量の把握,” 情報通信統計データベース, August, 2013.
- [4] Jon Postel, Ed., “Transmission Control Protocol, ” DARPA Internet Program Protol Specification, IETF Internet Standard, RFC 793, September 1981.
- [5] Jon Postel, Ed., “Internet Protocol, ” DARPA Internet Program Protol Specification, IETF Internet Standard, RFC 791, September 1981.
- [6] Ning Wang, et al., “An overview of routing optimization for internet traffic engineering, ” IEEE Communications Surveys and Tutorials, Volume 10, Issue 1, Page 36-56, April 2008.
- [7] Standardization, I.O.F., ”ISO/IEC 7498-1: 1994 Information Technology-Open Systems Interconnection-Basic Reference Model: The Basic Model,” International Standard ISOIEC 74981, March 1996.

---

## Chapter 2

---

### トラヒックエンジニアリング関連技術

ここまでインターネットの利用形態の変化と、通信に対する要求の変化、そして TE の必要性について述べた。本章では、この TE をインターネット上で行う上での課題や技術背景、そして関連する研究について述べる。

既に述べた通り TE の本質は、ネットワーク上で交換されるデータトラヒックを巧みに操作し輻輳のない効率的なネットワークを実現する事である。そして既存のインターネット技術に対して TE を実現する上で様々な課題を持ち、これらを本論文で扱う。第 1 章で述べた TE を実現する各レイヤーに関連する要素技術を図 2.1 に示した。本章ではこの図に沿ってこれらの技術と、関連する研究について述べると共に、本論文における各研究の立場を明らかにする。ここで、図中の各要素技術に関して簡潔に述べる。

図の一番下に位置する伝送プレーンは、パケットを処理する端末から次の端末までデータを送信するための技術である。このレイヤーの技術によってリンクの帯域が決定される。より細かく区分した場合、物理的に信号を表現して送信する技術と、データの送信および受信を制御する技術の二つに分類される。前者を物理層、後者を MAC(Media Access Control: 媒体アクセス制御)層と呼ぶ。物理的な信号としては主に光もしくは電気を用いることが一般的であり、またコアネットワーク以外であれば電波という選択肢もある。また MAC 層技術には、端末間で同期が行われているか否かが大きな違いとなる。実際には、多くの技術はこの両層にまたがって

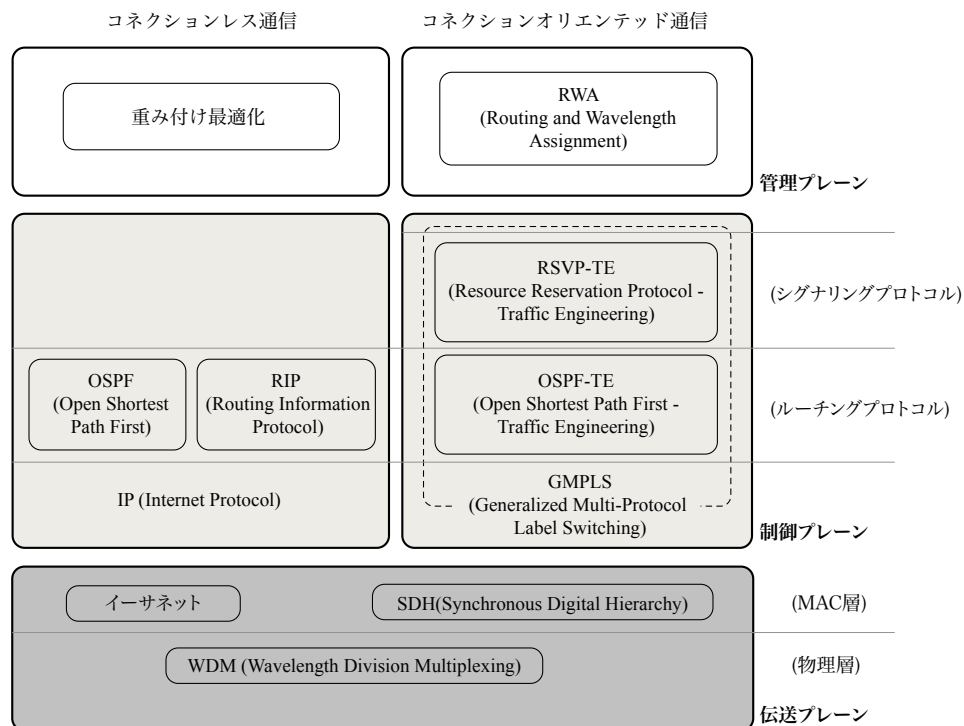


図 2.1: TE を実現の 3 つのレイヤーの各技術

おり，切り分けが非常に難しい。

SDH(Synchronous Digital Hierarchy) は同期型の通信方式であり，インターネットのようなパケット通信ネットワークに限らず遠距離通信技術として使用されてきた。SDH では STM-1 と呼ばれる 155.52Mbps の論理回線を単位として TDM(Time Division Multiplexing) による多重化を実現する。これによりネットワークを構成する事が可能だが，基本的に経路は固定的か，もしくは制御プレーン技術に任せる事となる。一方のイーサネットは，LAN(Local Area Network) における通信技術として発達した。1本の通信バスを複数の端末で共有する技術から発達した非同期型の通信である。現在ではブリッジングと呼ばれる機能により，イーサネットだけで小

規模なネットワークを構築することも可能だが、その仕組み上、閉回路を持つことが許されず、またルーチングのような高度な制御を行うことができない。そのため、コアネットワーク技術としては他の制御レイヤ技術を必要とする。一方の物理層の技術である WDM(Wavelength Division Multiplexing) は、1つのファイバー上に波長の異なる複数の信号を通す事により、ファイバーあたりの通信帯域を広げる技術である。特に通信帯域が必要とされるバックボーンネットワークにおいては欠かすことができない。このイーサネットおよび WDM については、本研究と関連が深いため、後に詳述される。

制御プレーンの技術は、パケットを処理する各ネットワーク端末が、次の端末を決定するための技術である。先に述べた通り、このレイヤーは通信方式によってコネクションレス通信方式の IP およびコネクションオリエンテッド方式の GMPLS に分類される（本論文では MPLS を GMPLS の一部として取り扱う）。また、管理プレーンについても、このそれぞれの通信方式によって技術が異なるため、両レイヤーは二つに分かれている。IP は宛先アドレスに基づいたコネクションレスの転送方式である。この方式では、基本的に同じ宛先アドレスを持つパケットは全て同じ経路を辿ることとなる。このような制約は、効率的な通信を実現する上で TE の柔軟性を損なう事となる。実際には、同じ宛先アドレスに対して複数の隣接ノードに等分配する ECMP(Equal Cost Multi Path) と呼ばれる技術もある。しかし、GMPLS のようなコネクションに基づく TE の柔軟性には及ばない。このような転送方式では、宛先アドレスに対する次のノードの決定に技術を要する。これを担うのがルーチングプロトコルであり、代表的なルーチングプロトコルに RIP( Routing Information Protocol)[1] と OSPF(Open Shortest Path First)[2] が挙げられる。OSPF ではメトリックと呼ばれる各ルーター間の距離をネットワーク管理者が自由に設定する事が可能である。この距離に大小を付けることを重み付けと言い、重み付けは各パケッ



トの経路を相対的に調節する。TEのための適切な重み付けは、管理プレーンにおける重要な技術である。

一方のGMPLSではコネクションに基づいたフローベースの転送を行う。フローとは送信元から宛先まで決められた経路を通過する一連のパケットの流れである。フローベースの転送により、各パケットは同じ宛先であっても別の経路を通過することも可能となる。これにより、より柔軟なTEを行う事が可能となる。各フローのためにネットワーク上に確立される固定経路をLSP(Label Switched Path)と呼ぶ。LSPの制御を担うのが、シグナリングプロトコルである。通信をLSPに基づいて行うGMPLSでは、LSPの制御に多く課題を持つ。特に、障害時にどのように通信を存続させるかは大きな課題である。G. Iannacconeら[3]によれば、バックボーンネットワークにおけるリンク障害はいまだ頻繁に発生しており、これらの課題には、障害時に備えたバックアップ経路の準備や、障害発生時の素早いLSPの切り替え手法が求められる。さらに、GMPLSによるTEは、単なる輻輳の解決だけではない。TEによる遅延やパケット損失などのQoS(Quality of Service : サービス品質)を保証したパス確立などがそれに相当する。既に述べたように、昨今のマルチメディアアプリケーションは帯域だけでなく、QoSに関しても保証を要求する。TEを用いてこれらの要求を満たすことは、ネットワークプロバイダにとって大きなチャレンジである。また一方で、マルチキャストはTEによって実現すべき課題である。通信系路上にデータを複製を含むこの通信では、ネットワーク上に複製の指示を保持しなければならない。コネクションレス型の通信方式では実現が困難であった。GMPLSのようなコネクションオリエンテッドなネットワークによるマルチキャストの実現が期待されている。

またGMPLSによるTEを行う上で、下位レイヤーである伝送プレーンとの互換性は大きな課題となる。先に述べたように、伝送プレーンの技術としてSDH、イーサ

ネット，WDMのネットワークなどでは，各技術に特化したTEが求められる．特に，近年では低コストな広帯域技術として注目を浴びているイーサネットは，GMPLSのTEとともにWANとしての新しい技術として注目をされている．

本章では，このTEを実現する上での関連する技術について述べる．最初にTEを実現するに至った経緯や，ネットワークの基本的なメカニズムについて前半で説明する．そして後半では，昨今のTEに関する研究の紹介と，本稿の位置づけを行う．

## 2.1 トラヒック制御のためのネットワーク基盤技術

### 2.1.1 制御プレーン技術：IP ネットワーク技術

インターネットにて用いられるIP(Internet Protocol：インターネットプロトコル)は，代表的なベストエフォートプロトコルであった．1970年から80年にかけて策定されたこのプロトコルは，電話のようなリアルタイムなアプリケーションではなく，コンピュータ間でのデータの転送のために発明された．インターネットでは，転送されるデータはパケットという単位に分けられ，各パケットはネットワーク上の各リンクを1ホップごとに転送されていく．データを送信するコンピュータから受信するコンピュータまでの間に，通常は複数のリンクがあり，ルーターによってパケットは中継される．ルーター（および送信するコンピュータ）は送ろうとしているパケットを，既に決定された経路上の隣接ルーター（もしくは受信するコンピュータ）へと送る事に責任を持つ．そのため，始点から終点までの接続性は保証されないが，その代わりとして非常に高いリソース共有率を発揮する．本来は一つのプロトコルであったTCPと共に，TCP/IPとして動作する事によって，ネットワーク全体でできる限り公平な通信を実現する．

IPにおいて，ルーターはパケットの宛先アドレスに対して決定づけられた隣接

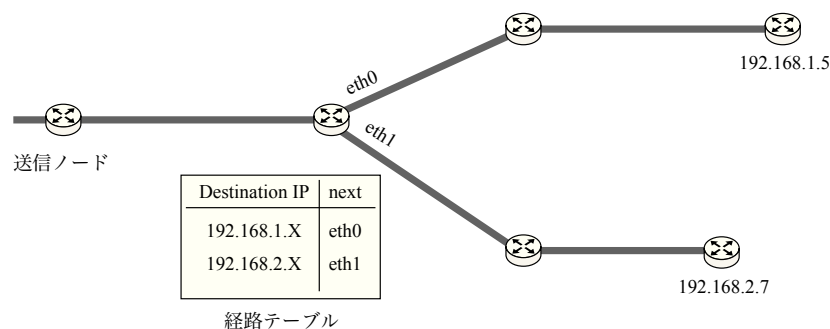


図 2.2: IP フォワーディング

ルーターへと送信する。このより近いルーターの方向は、予めルーター内に経路テーブルとして記されている。パケットを受け取る度にこの経路テーブルを確認し、次の隣接ルーターを決定するこの仕組みを IP フォワーディングと呼ぶ。図 2.2 へと IP フォワーディングの例を示す。この経路テーブルはパケットが到着している時点で構築していなければならない。経路テーブルの構築をルーチングと呼び、ルーチングプロトコルはルーチング自動化のためのプロトコルである。

もっとも歴史のあるルーチングプロトコルである RIP は、ディスタンス・ベクタ・アルゴリズムに基づいてルーチングテーブルを構築する。ルーターは自分のネットワークアドレスを含むパケットを隣接ルーターに送信し、各隣接ルーターはメトリックと呼ばれる距離を示す値を加算してそのパケットを隣へ転送する。最終的に、ネットワーク上の全てのルーターは対象となるネットワークアドレスまでの最短経路を把握する事ができる。RIP によるルーチングは、非常にシンプルな実装で最短経路を導出する。しかし、パケットの到達可能性が主な課題であった当時の技術背景から、RIP はリンクの距離という概念を持っていなかった。そのため、メトリックの値は 15 を最大値とし、1 ホップごとに 1 を加えるという使い方であった。その

他にもループを発生する危険性や、トポロジー変化に対する情報の収束の遅さなどが問題とされ、ネットワーク機器の性能向上とともに OSPF に代わられて行った。

#### 2.1.1.1 OSPF プロトコル

OSPF は現在、最も一般的に使用されるドメイン内ルーチングプロトコルである [4]。RIP がネットワーク上で最短距離情報を交換するという基本概念に基づくのに対し、OSPF は各ルーターの接続情報をネットワーク全体に公開するという考えに基づいている。言い換えると、RIP は「それぞれの宛先に対してどちらが近いか」だけの情報を交換するのにに対し、OSPF では「どのノードとどのノードが接続されているか」の情報を交換する。これによって、全てのノードはネットワーク全体の構造を知る事ができる。さらに、OSPF ではメトリックはホップ数ではなく、ノード間の仮想的な距離としてネットワーク管理者によって設定される値となる。メトリックを含むノード間の接続情報をリンクステートと呼び、OSPF はリンクステートルーチングプロトコルと呼ばれる。図 2.3 にリンクステート交換の例を示す。

図のように、ノード A が発行するメッセージは、ノード A に関する隣接ノードの情報を含む。この発行されたメッセージは、フラッディングによりネットワーク全体へと配られる。ネットワーク全体への情報配布を、OSPF ではリンクステートの広告と呼ぶ。効率的に、且つメッセージのループを起こさずにフラッディングによる広告を実現することが OSPF では求められており、そのため実装は RIP よりも複雑となる。また、それぞれのノードの隣接情報が一つのメッセージを発行し、そのメッセージが最終的に全てのノードへと到着することを考えると、各ネットワーク端末への負荷は大きい。しかしながら、定期的な情報交換に基づく RIP に比べ、より大きなドメインのネットワークにおいて利用可能であり、通信機器の向上も背景に、OSPF は広く使われてきている。

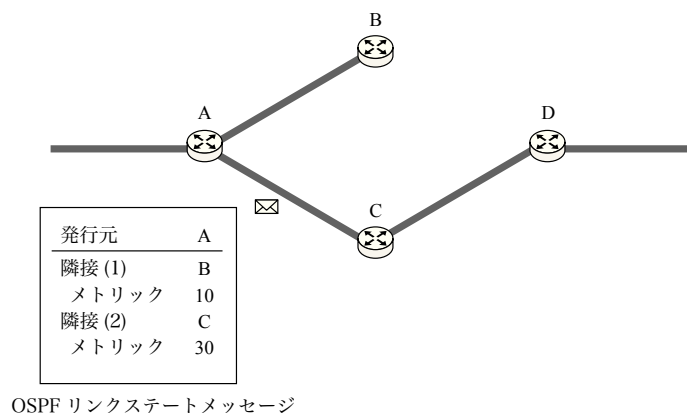


図 2.3: OSPF によるリンクステートの交換

これらの概念を実現するための実際の OSPF プロトコルの実装はもう少し複雑である。ノード間のハンドシェイクの方法や、スタートアップ時の挙動、トポロジー変更時の挙動なども厳密に定義されている。図 2.4 に実際の広告メッセージのフォーマットを示した。リンクステートが変更した際に、古い情報か否かを調べる為の LS Sequence Number(リンクステートシーケンス番号) や、情報が発行されてからの時間を示す LS Age なども含まれる。ここで、トラヒックエンジニアリングを実現する上で最も重要なのがメトリック (Metric) フィールドである。メトリックの最大値が 15 である RIP ではホップ数をメトリックとして計算するのに対し、2byte の値で 1 から 65536 までの値が許される OSPF では、メトリックはネットワーク管理者によって予め自由に設定できる値である。このメトリックはノード間の仮想的な距離であり、各ノードは合計値が最短となる経路を計算して、経路テーブルに記入する。この値は重み付け方がトラヒックの経路に作用する事から、OSPF 重み (OSPF Weight) とも呼ばれる。例えば、大きなメトリック値を持つリンクは、最短経路計

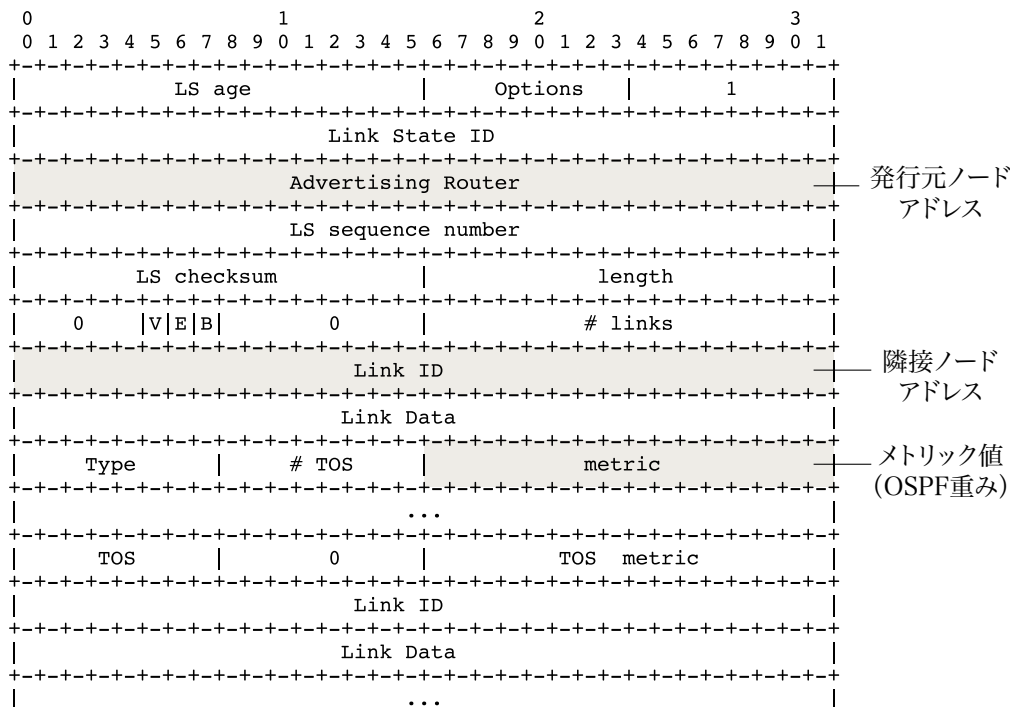


図 2.4: OSPF メッセージの例

算において経路に選ばれにくくなり、相対的にトラヒックが流れる確率が低くなる。また、メトリックの値を消費電力などとすることによって、電力消費を最小とするような TE も可能となる。最短経路導出の手法と、このメトリックを使用した負荷分散の方法については次に説明する。

### 2.1.1.2 経路計算と ECMP

OSPF のネットワークでは、フラッディングによって各隣接ノードとの接続情報をネットワーク全体に配布する。今度は、この配布が全て完了した後に最短経路を計算し、経路テーブルを構築するプロセスについて図 2.5 を用いて説明する。

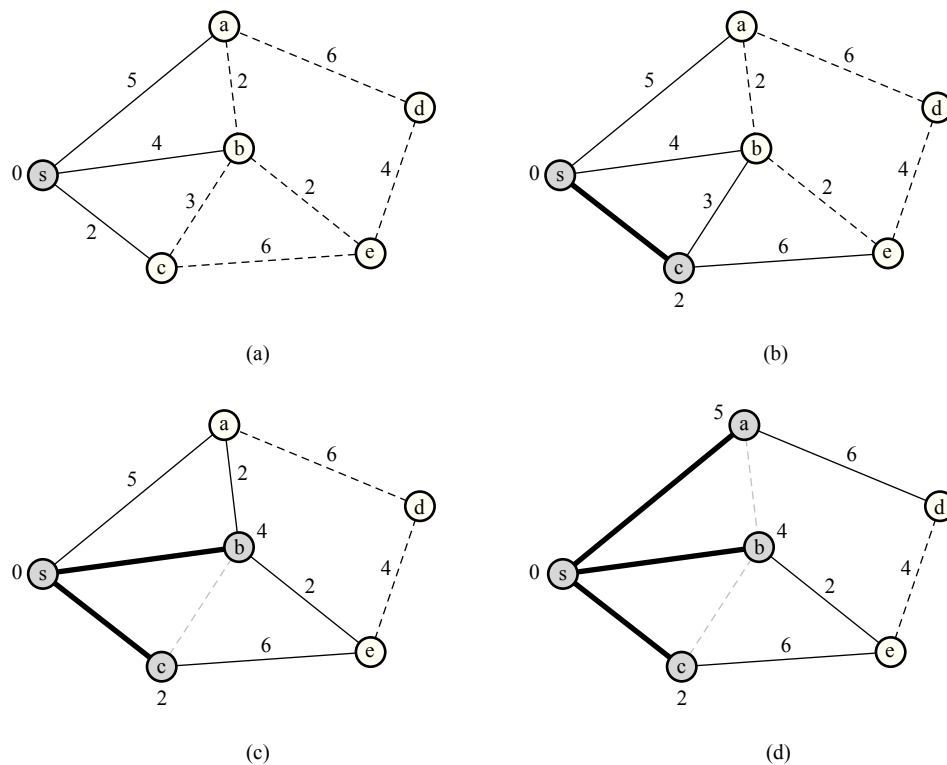


図 2.5: ダイクストラ法

図 2.5 は最短経路探索アルゴリズムの一つであるダイクストラ法の例である。いま、ネットワーク上のノード  $s$  が経路テーブル構築のために、各ノードへの最短経路の計算を行っている。ダイクストラ法では、最短経路と距離が確定したノードから、その隣接ノードの最短経路を決定してゆく。まず最初に、自身への距離は 0 であることを確定する。確定したノードは図中 (a) のようにグレーに塗りつぶす。そして次に、最短距離が確定した全てのノードの隣接するノード（ここでは候補ノードと呼ぶ）の距離を計算する。図の例ではノード a, b, c が候補ノードとなる。候補ノードの中で、最も距離が小さいノードは、他に最短距離を持たない事が明白である。よっ

てノード  $c$  が確定され (図中 (b)), 次の候補ノードは  $a, b, e$  となる. これを繰り返す事によって, 一つずつノードが確定してゆく. 最終的には,  $s$  を根とするツリー構造を構築する. このツリー構造から, 隣接ノードを決定し, 経路テーブルを構築する事が可能となる.

ここで, 図中 (d) の, ノード  $a-b$  間のリンクや  $b-c$  間のリンクは, ツリーに含まれない. これは即ち, ノード  $s$  を出発するメッセージは全てこのリンクを使用しない事を意味する. もしもネットワークの全てのトラフィックがノード  $s$  からしか発生しないとすれば, 非効率な通信となる可能性は明白である.

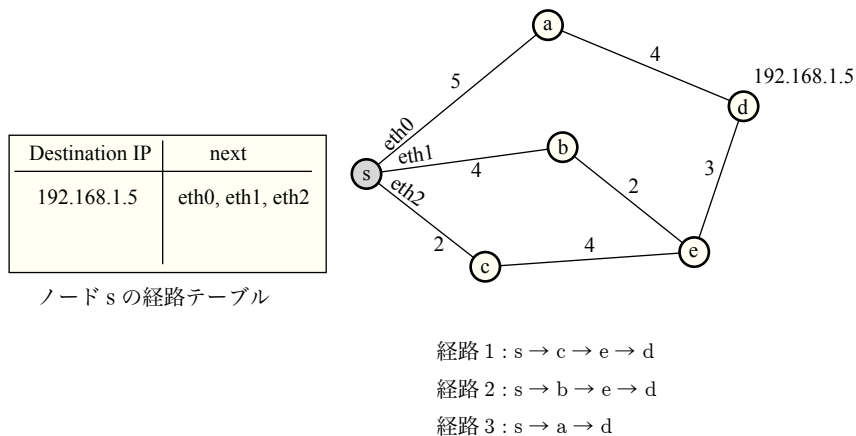


図 2.6: ECMP

一方で, OSPF は最短経路ルーティングにおいても, 最低限の負荷分散の機能をサポートする. これは, 距離が同一である経路が複数あった場合に, 経路テーブルに次ノードを複数記述することによって, 同じ宛先のトラフィックを複数の経路で送信する技術である. 図 2.6 に ECMP の例を示す. ここでは  $s$  から  $d$  までに 3 つの経路が存在する. この場合, トラフィックを 3 つに等分して送信することによって負荷



分散が実現される。この図では経路1および2は同じリンク  $e-d$  を通過するため、3等分が最適とも言えないが、ECMPは従来の一つの経路しか通過しない送信方法よりも遥かに高い通信効率を実現する。これらの詳細については、後にIPベースのTEとして紹介する。

### 2.1.2 制御プレーン技術：GMPLS ネットワーク技術

ここまで、IPによるパケット転送の仕組みと、経路テーブルを作成するための二つのルーティングプロトコル、RIPとOSPFを紹介した。OSPFの登場により、後に説明されるような各リンクのメトリックの重み付けによるTEが実現される。しかしながら、宛先アドレスに基づいたIPの通信方式では、経路を自由に制御する事はできない。そこで、コネクションに基づいた通信によって経路を自由に制御し、柔軟にTEを実現するMPLSが開発された。そして、GMPLSはMPLS一般化拡張したものであり、さまざまなネットワーク上にLSPと呼ばれる固定的な経路を確立することによって、トラヒックの流れを自由に制御する。GMPLSはその前身であるMPLSの拡張として提案され、現在も標準化の段階にある。本論文では、MPLSはGMPLSの一部として扱う。まずは、MPLSおよびGMPLSの通信の仕組みについて簡単に説明する。

#### 2.1.2.1 MPLS ネットワーク

MPLSの当初の目的は、パケットが到着する度に行われる次ホップの決定の高速化であった。インターネット全体を網羅するIPネットワークではアドレス集約の利便性から、プレフィックスを用いた経路テーブルを許可した。図2.2では、宛先ノードアドレスである192.168.1.5の変わりに、経路テーブルにはプレフィックスと

して 192.168.1.X だけが記載されている。このプレフィックスを用いた集約による記載法が、IP 空間全体に対する経路を記述する単純化する事は明白である。しかしながら、プレフィックスを用いた経路表におけるアドレスの探索は、最長一致探索を必要とするために処理時間に時間がかかり、これがパケットの処理遅延につながっていた。この経路テーブルの探索を完全一致探索とすることにより処理時間を短縮する事が、MPLS の当初の狙いであった。ラベルと呼ばれる識別子による完全一致探索のために、コネクションに基づいた転送を行うのが MPLS である。実際には、TCAM (三値連想メモリ：Ternary Content Addressable Memory) の開発によって最長一致も固定時間での処理が可能となったため、IP 転送に比べ探索が高速となる優位性は低く、MPLS の利点はコネクションに基づく高度な TE を有するパケット転送技術となる。

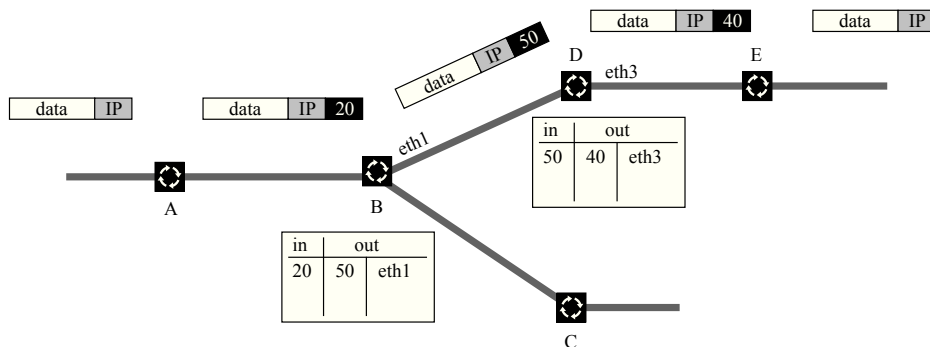


図 2.7: MPLS のラベル転送

図 2.7 に MPLS におけるラベルに基づくパケット転送方式を示す。各パケットに対する経路の決定は、MPLS ヘッダ (Shim ヘッダと呼ばれる) に含まれる 20 ビットのラベルと呼ばれる値によって行われ、この転送方式をラベルスイッチングと呼ぶ。

ラベルの値は各ノードで変更されるため（この処理をスワップと呼ぶ），他のノードにおいてラベル値の衝突は発生しない．入力ラベルに対し，出力ポートと出力ラベルが一意に定まっており，これにより一連の固定的なフローが形成される．図中においても，ノード A をラベル 20 で出発したパケットは，ノード E へとラベル 40 で到着する．このような同じフローを構成するパケットの集まりを FEC(Forwarding Equivalence Class：転送等価クラス)と呼ぶ．そして，始点ノードから終点ノードまで構成された経路を LSP(Label Switched Path)と呼ぶ．

このラベルの関係性は，IP における宛先アドレスによる転送とさほど変わらない．IP ネットワークにおいてルーティングプロトコルによって経路テーブルが作成されるのに対し，MPLS ではルーティングプロトコルとシグナリングプロトコルによって経路テーブルが作成される．このメカニズムについては後に説明される．また，多くの場合 MPLS では IP パケットをカプセルリングして転送する．図中では，ノード A に到着した IP パケットが，ノード E までは MPLS ネットワークとして転送され，ノード E からは IP として転送されている．ここで，MPLS において始点となるノードを始点ノード，終点となるノードを終点ノードと呼び，パケットの送信元ノードおよび宛先ノードとは分けて呼ぶ事とする．

MPLS におけるコネクションに基づいた転送方式は，本論文で扱う TE を実現するために非常に適していた．ラベルの転送テーブルさえ決めてしまえば，LSP によって任意の経路にトラヒックを通す事ができる．この転送方式を様々な伝送プレーン上で行うものが GMPLS である．

### 2.1.2.2 GMPLS ネットワーク

GMPLS はあらゆるネットワーク上でコネクションに基づいた通信を実現するためのフレームワークであり，MPLS の拡張である．MPLS のラベルの考えを一般

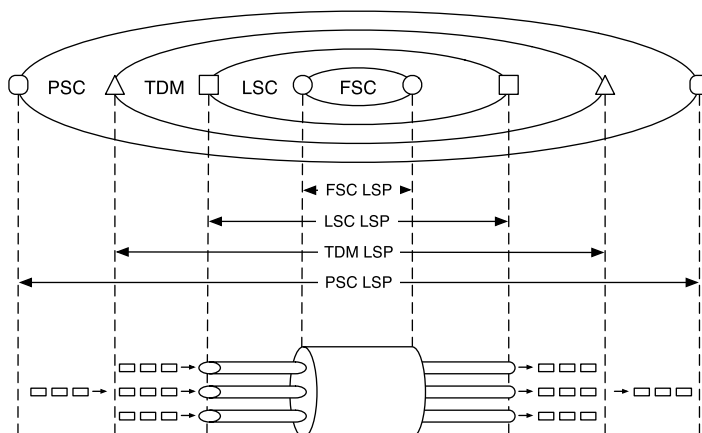


図 2.8: GMPLS のレイヤー

化し、さまざまなネットワーク上で利用することができる。現在は IETF によって  
 プロトコル詳細が標準化されている段階である。

MPLS では MPLS ヘッダに含まれる 20bit のラベルを用いてパケットの経路を  
 決定した。その値はスワップにより経路上で変化するが、ラベルは LSP の識別子で  
 あると言える。この識別子に、Shim ヘッダ内の値だけでなく、既存の様々なネット  
 ワークの特性を利用したのが GMPLS の一般化ラベル (Generalized Label) である。  
 一般化ラベルの概念は (1)TDM ネットワークにおける、スロットの位置、(2)WDM  
 ネットワークにおける波長、(3)トラヒックが通過中のファイバーなどに拡張され  
 る。これらはヘッダ内に示された識別子ではないが、それぞれに基づいて経路を決  
 定する事ができる。GMPLS では、上記のそれぞれのスイッチングを 4 つのタイプ  
 分類し、スイッチングケーパビリティ (Switching Capability) と呼ぶ。各スイッチ  
 ングケーパビリティの関係は図 2.8 に表される。

- PSC(Packet Switching Capability : パケットスイッチング) : パケット内の

ヘッダ情報に基づいたスイッチングである。MPLSにて使用されているスイッチングもここに分類される。ATMネットワークもGMPLSではサポートされている。

- TDM(Time Division Multiplexing : 時分割多重) : TDMネットワークにおいて、データがどの位置に存在するかをラベルとして用いたスイッチングである。主にSONET / SDHに関して用いられる。
- LSC(Lambda Switching Capability : 波長スイッチング) : WDMネットワークにおける信号の波長に基づいたスイッチングである。OXC(Optical Cross Connect : 光クロスコネクト)等の装置で、波長に基づいて経路を決定するためのものである。
- FSC(Fiber Switching Capability : ファイバースwitching) : ファイバー単位のスイッチングに用いられる

GMPLSはTEの考え方を大幅に広げ、様々なレイヤーにおいてトラヒックの制御を可能とした。特に、WDMネットワークなどではパケット毎のスイッチングは困難であり、固定的な経路にトラヒックを通過させなければならない。これによって、IPにとって1つのリンクを形成するだけでしかない各伝送プレーン技術は、GMPLSにおいてはネットワーク技術として利用可能となる。

### 2.1.2.3 OSPF-TE プロトコル

GMPLSを支えるルーティングプロトコルの一つが、OSPF-TE(Open Shortest Path First - Traffic Engineering)である。OSPF-TEはOSPFのTE拡張であり、基本的な動作はOSPFと同じである。

しかしながら、OSPF-TEではGMPLSがサポートする様々なスイッチングケーパビリティに関するリンクステートを広告する。これらのリンクの多様化のために、OSPF-TEはリンクステートに関する拡張の余地を残している。もっとも大きな違いは、GMPLSのためのOSPF拡張[5]でサポートされるリンク帯域の概念である。IPネットワークにおけるルーチングプロトコルのRIPやOSPFでは、あくまで単一のメトリックという値を経路選択に使用していた。またトラヒックエンジニアリングの概念を持たないこれらのネットワークでは、ベストエフォートにより到達性だけを考慮したパケット転送が行われる。GMPLSではフローに対して厳密に通信帯域を定義し、これを管理する事ができる。実際には通信帯域は、Maximum Bandwidth(リンク最大帯域) / Maximum Reservable Bandwidth(最大予約可能帯域) / Unreserved Bandwidth(未使用帯域)に分けられる。そしてこのUnreserved Bandwidthに関しては8つの優先度が定義される。その他にも、遅延、損失率の値や、物理レイヤーに関係するパラメータをメトリックとして扱うOSPF-TE拡張も提案されている。これらの詳細に関しては、後の関連研究の項にて紹介する。

#### 2.1.2.4 RSVP-TE プロトコル

OSPF-TEによって配布された情報に基づき、経路が決定されると、今度はその経路上にLSPが確立される。GMPLSにおけるLSPの確立はRSVP[6]のTE拡張であるRSVP-TE(Resource Reservation Protocol - Traffic Engineering)[7]とよばれるシグナリングプロトコルによって行われる。

始点ノードで経路が決定された後、RSVP-TEのシグナリングによってLSPが確立される。そして確立したLSPを制御するのも同じようにRSVP-TEが担当する。RSVP-TEにおける最も大きな課題は、どのようにしてリソース予約の競合なく資源を予約し、そしてリンク障害などの不足の事態を適切に対処するかである。

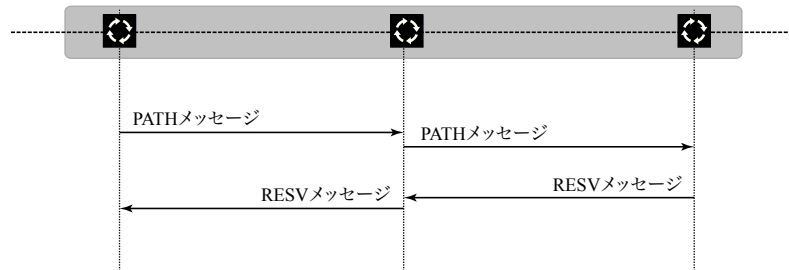


図 2.9: RSVP-TE によるシグナリング

RSVP-TE では、まず確立しようとしている LSP の経路上をメッセージが通過し、その時にネットワークリソースが確保される。図 2.9 に RSVP-TE によるシグナリングの例を示す。始点ノードから終点ノードへと送られるメッセージを PATH メッセージと呼び、このメッセージには確立している LSP の種類、そして必要な帯域、また終点ノードまでの明示的な経路が含まれる。PATH メッセージが終点ノードへと到着すると、終点ノードは確認応答のためのメッセージを始点ノードへと送る。これを RESV メッセージと呼ぶ。もしもメッセージ到着時に十分なリンク帯域がない場合はエラーとなり、ERROR メッセージを始点ノードへと戻す。通常、経路や帯域の情報は OSPF-TE によって各ノードへと予め配布されており、始点ノードはその情報に基づいて経路を決定しているため、リソース不足による予約エラーは稀である。一方のラベルは、OSPF-TE で配布を行わないため、この RSVP-TE によるシグナリングにおいて決定される。MPLS では、各ルーターは受信したパケットのラベルの値によって、そのパケットに対する操作（ラベルのスワップおよび転送）が決定される。これはすなわち、下流側のルーター（すなわち、終点ノードに近い側のルーター）がラベルを管理するのが適切である。このポリシーに従って、RSVP-TE では RESV メッセージにおいてラベルが決定される。ここで、WDM ネットワーク

においては、ラベルの値が波長である。波長コンバータなどの制約から、上流側にも使用できるラベルに制限が含まれる。そのため、RSVP-TEはラベルセットオブジェクトによって上流側から使用できるラベルを提示する機能も含まれる。

ラベルの決定がRESVメッセージに含まれることから、通常スイッチングテーブルの変更はRESVメッセージを処理する際に行われる。RESVメッセージが変更され、スイッチングテーブルが変更されることによって、LSPが確立されていく。ここで、RESVメッセージが途中までリソースを予約し、スイッチングテーブルを変更した後に、エラーを起こしたとする。その場合、エラーを起こしたノードは途中まで確立したLSPを切断しなければならない。そのためにTEARメッセージを下流へと送信する。

また、リンクに障害が発生した場合の対処も、RSVP-TEの大きな役割である。既に説明したように、IPネットワークではホップ毎に経路が決定されるのに対し、GMPLSではLSPの確立によって、自由に経路が制御される。

ここで、ネットワークにおいてリンク切断などの障害が発生した場合を考える。IPネットワークでは、障害が発生した場合、各ノードで経路の再計算が行われることにより復旧される。障害が発生したリンクに近いノードから順に障害に関する情報を受け取り、再計算が行われ、各ノードの経路テーブルが書き換えられる。経路テーブルが書き換えられるまでの間は、通信は行われない。一方のGMPLSでは、リンクの障害によってリンク上を通過していた全てのLSPが切断される。そして切断されたLSPの始点ノードが、新たにLSPを確立することによって障害に対処する必要がある。リンクの障害をいち早く察知する為に、RSVP-TEではリフレッシュメッセージを定期的に交換する。リフレッシュメッセージはLSPの始点ノードから終点ノードまで定期的に往復し、リンク障害などによるLSP切断を検出する。その後、始点ノードは迂回ルートのための経路再計算を行い、シグナリングを行って別



の経路上に LSP を確立する。これを仕組みをリストレーションと呼ぶ。リストレーションでは、障害が発生したリンク上を通過する全ての LSP に対して再セットアップを必要とする。これらの再セットアップのためのシグナリングは、互いに衝突する可能性があり、結果として収束するまで通信不能となる場合がある。コアネットワークにおいては、そのような通信障害はあってはならない。特に、通信帯域や効率性よりも、耐障害性を重要視するようなネットワークではリストレーションによる復旧は信頼性に欠けると言えるだろう。

そこで、予め各リンクが切断した際の迂回経路を予め設定するプロテクションと呼ばれる方式が提案された [8]。このプロテクションでは、バックアップ LSP を予め確立しておく事により、障害発生時に瞬時に通信の復帰を実現する。特に、一つのリンク障害に注目した際に、効率的にバックアップ経路を選択するなど、様々な研究が行われた。これらの研究は Marzo Jose ら [9] によってサーベイされている。

### 2.1.3 管理プレーン技術：トラヒックのモデリングと

#### トラヒックエンジニアリングの実現

ここまで、各ネットワークにおいてパケットの経路が決定されされるまでのメカニズムについて説明した。コネクションレス通信である IP ネットワークと、トラヒック経路を決定する OSPF、そしてコネクションに基づいたフレキシブルな TE 通信を実現する GMPLS について述べた。今度は、ネットワークが扱うトラヒックについて注目をする。

図 2.10 の (a) に表したのは、ネットワークをブラックボックスとして見た図である。ネットワーク上の  $a, b, c, d, e$  の 5 つのノードからパケットは送信され、このいずれかのノードへと到着する。ネットワーク内部でどのように転送が行われるにし

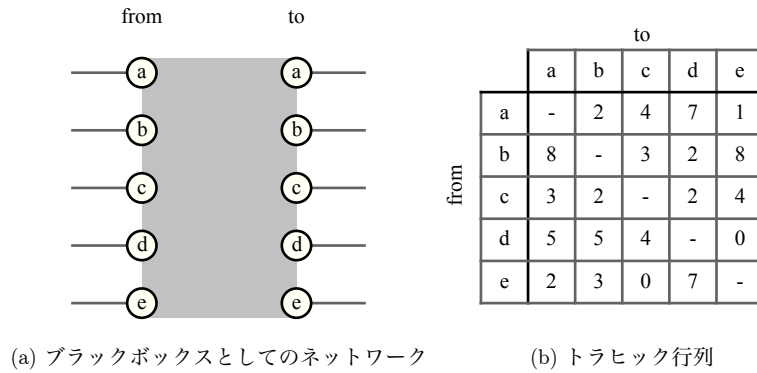


図 2.10: 抽象化したネットワークの通信モデルとトラフィック行列

ても、ネットワークに対する通信の要求はこのモデルに帰着する。ここで、ある一定の時間内に交換される、もしくは交換されたトラフィック量に注目する。これらは図中 (b) のように送信元と宛先の行列として表される。この行列をトラフィック行列という。通常、この行列内の数字は単位時間あたりのデータ量 (Mbps や Gbps など) である。

このトラフィック行列が何らかの形で与えられた場合に、ネットワーク上の輻輳を回避するというのが TE の一つの方法である。トラフィック行列を推定する方法は、大きく分けて二つある。一つは SLA (Service Level Agreement : カスタマーサービス契約) に基づく方法である。SLA はカスタマーと ISP の間で交わされた通信サービスに関する詳しい仕様書であり、多くの場合通信する帯域について記載されている。すなわち契約書に記された通信帯域の上限値である。多くの場合、この上限値を超えたトラフィックがネットワーク内へと進入しないよう、境界となるルーターにてトラフィックシェーピングと呼ばれるトラフィック流量の調整を行う。もう一つの方法はモニタリングである。モニタリングの手法に関しては [10] が挙げられる。これらの方

法からトラヒック行列を推定し、最適な経路割当を計算し、新たな設定を適用する。このトラヒック行列の推定と、経路計算、そして設定の適用を反復して実行するのが、オフライン TE である。オフライン TE において重要な問題は、RPC(Resource Provisioning Cycle)[11] と呼ばれる TE を実行する時間間隔である。図 2.11 にオフライン TE のモデルと RPC の関係を示した。多くのオフライン TE では、RPC は一週間から一ヶ月程度である。その場合、バーストトラヒックやネットワーク障害などの変化に動的に対応した TE は難しい。推定したトラヒック行列からのずれは、TE を非効率化する。

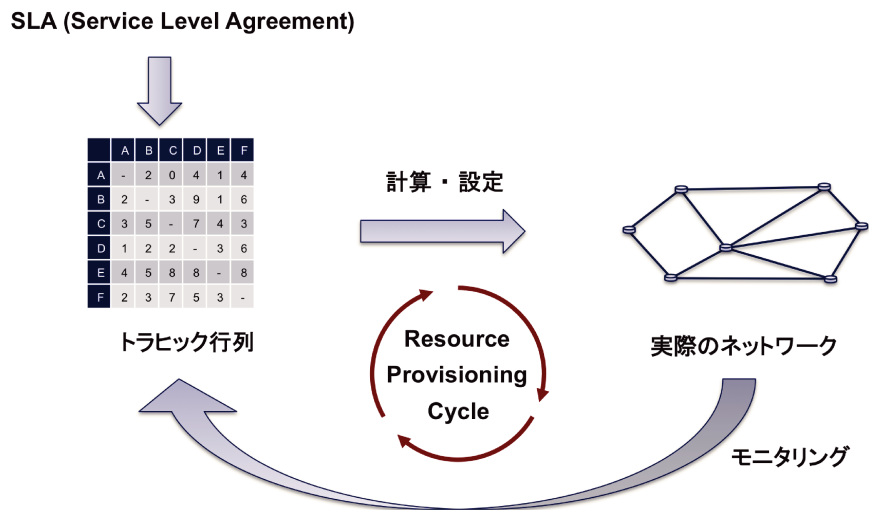


図 2.11: オフライン TE と RPC

IP における TE は、OSPF メトリックの値を調節する事によって、経路を制御し負荷分散を実現する。トラヒック行列が既に与えられていた場合、ネットワーク全体の混雑度を最も低くするメトリックが計算できる。この問題を OSPF 重み付け問題と呼び、OSPF 重み付け問題は典型的なオフライン TE である。これらについ

ては本章の関連研究の項および第3章にてより詳しく説明される。

一方で、トラヒック行列が求められない場合や、またフロー単位での通信が要求される場合などでは、トラヒック行列を用いないオンラインTEが行われる。オンラインTEの主たる課題は、どのように動的なシステムに自己収束を実現させるかである。到着したトラヒック要求を処理する際に、それ以降に到来するトラヒック要求によって輻輳が起これないように、最適に配置することである。既に配置したトラヒックフローの再配置を伴う事も可能であるが、この再配置はネットワークを不安定としたり、サービスの一時的な停止を発生させるほどの規模で行ってはならない。近年では、リソース管理の単純化の目的から、ネットワーク上のいずれかに配置したPCE(Path Computation Element：経路計算装置)[13]によって集中的に計算する手法も提案されている。しかしながら、集中管理の手法は、分散処理による耐障害性を損なうという問題もある。

#### 2.1.4 管理プレーン技術：マルチメトリック環境での経路選択

RIPにおけるルーチングでは宛先までに通過するルーターの数、即ちホップ数を基準として最短経路が選択された。このホップ数をメトリックと呼んでいたのに対し、OSPFではメトリックは管理者によりリンクに割り当てられるネットワーク上の仮想的な距離である。メトリックは言葉通り、リンクの「尺度」であり、計り方によって距離や重みが変わるものである。そして、GMPLSネットワークでは、TEのために「リンク帯域」という新たな尺度をメトリックとしてOSPF-TEに導入した。OSPF-TEではさらに遅延、損失率や、物理レイヤーにおけるパラメータなど様々なリンク特性をメトリックとして配布する為の拡張を許している。このように、様々な値をメトリックとして経路の選択を行う環境をマルチメトリック環境と呼ぶ。

OSPFにおけるルーチングでは、ダイクストラ法によってメトリックの合計値

を最小とする最短経路を導出した。このダイクストラ法は一般的に、単一のメトリックに対する最短経路導出に対して有効である。即ち、ノード  $i$  から  $j$  にへのリンクにおけるメトリックを  $d(i, j)$  とした場合に、ダイクストラ法は

$$\text{minimize } d(p) = d(i, j) + d(j, k) + \dots + d(l, m)$$

となるような経路  $p = (i, j, k, \dots, l, m)$  を導出する。一方で、複数のメトリックの条件において、何を最小化すべきかについては議論が残る。多くの場合、全てのメトリックが同時に最小となるとは限らない。そのため、予めメトリックに優先度を付け、これを関数として式で表す必要がある。例えば  $i, j$  間の帯域を  $B(i, j)$ 、遅延を  $D(i, j)$ 、パケット損失率を  $L(i, j)$  とした場合、

$$f(i, j) = \frac{B(i, j)}{D(i, j) \times L(i, j)}$$

のような関数が考えられる。 $f(i, j)$  のような統合メトリックを定義する事により、マルチメトリック環境における経路計算がダイクストラ法によって行われる。

ここで、GMPLS における TE のように LSP へと固定的な帯域が割り振られる場合、帯域については各リンク帯域が要求される帯域を上回っていれば十分である。すなわち、帯域に関しては LSP 上の全てのリンクの最小値、ボトルネックとなるリンクの帯域だけが有効である。このボトルネックの帯域が、要求帯域を下回る場合には LSP を確立する事ができない、すなわち制約条件としての働くメトリックである。ここで、経路全体の制約条件としてのメトリックは加算的 (Additive)、乗算的 (Multiplicative)、そして最小値 (Concave) に3種類に大別される。先ほど同様に、LSP の経路を  $p = (i, j, k, \dots, l, m)$  と置いた場合に、 $d(p)$  が加算的とは

$$d(p) = d(i, j) + d(j, k) + \dots + d(l, m)$$

また、 $d(p)$  が乗算的とは

$$d(p) = d(i, j) \times d(j, k) \times \dots \times d(l, m)$$

そして、 $d(p)$  が最小値とは

$$d(p) = \min[d(i, j), d(j, k), \dots, d(l, m)]$$

ということを指す。それぞれの例として、遅延、パケット損失率、帯域などが挙げられる。これらの制約は、同時に存在する場合もあり得る。そして最小値に関する制約のみ、複数の制約が同時に存在しても、ダイクストラ法を拡張する事によって計算可能である。この拡張した制約下の最短経路計算を CSPF (Constraint-based Shortest Path First) と呼び、本章で後に紹介される。一方で、加算的もしくは乗算的およびこれらの組み合わせによる、2つ以上の制約があった場合、その解が存在するか否かの決定問題は NP 完全であると証明された [12]。即ち、解を求めるための計算が NP 困難となり、計算量が非常に大きくなってしまう可能性がある。実際には、多くのメトリックはホップ数に従って大きくなるため、最適解を求めないヒューリスティックな解法によって、十分に良い解を導く事ができる。しかしながら、その経路が制約を満たすか否かを確認する為に、各メトリックの合計値の計算は最低限必要となる。このように、マルチメトリックのサポートは、ネットワークの経路計算やリソース管理における課題である。

### 2.1.5 伝送プレーン技術：WDM

今度は下位レイヤーの技術へと注目する。OSI の階層モデルでは本来、上位レイヤーは下位レイヤーを意識せずに通信を行うという設計思想で開発された。即ち、経路を決定するネットワークレイヤーにとって、下位レイヤーの通信技術は単にパケットを隣接ノードへ伝達する 1 ホップのリンクにすぎない。しかし通信の品質を考えた場合に、そのリンクがどれほどの最大帯域を持つか、どれほどの遅延を生じるかなどは下位レイヤー技術に依存する。そのため、TE を行う上で下位レイヤー

の考慮が重要となる。下位レイヤーの技術で、特に TE を必要とする物理レイヤーの伝送技術の一つが WDM である。

WDM は同一の光ファイバー上で波長の異なる複数の光信号を用い通信を行う技術で、一つの波長が最大でおよそ 40Gbps の通信帯域での通信を実現し、また一つのファイバーに 64 から 160 個の波長が収容される。結果として、ファイバーあたりの通信帯域は Tbps 単位まで達する。これらの光信号をネットワーク上でスイッチングする場合、低速である電気信号への変換は得策ではない。そのため、光クロスコネクタなどの装置を用いて、光信号のままスイッチングされるが、高速なスイッチングや光デバイスが未成熟なことから、ホップ毎の経路選択は現段階では不可能である。そのため、WDM では GMPLS のようにコネクションに基づいた通信が必要となる。WDM ネットワークにおける LSP は、光信号を発信する始点ノードから光信号を終端する終点ノードまでの一連の経路であり、光パスと呼ばれる。

WDM における TE は、QoS や輻輳ではなく、物理層の制約に対して効率的な通信を行う為に用いられる。電気信号と異なり、光パスのスイッチングは信号に対して一切の処理を行わない。そのため、信号は始点ノードから受信ノードまでの間に信号損失や劣化を受け、場合によっては通信が行われなくなる可能性がある。GMPLS ではこれらの物理レイヤー上の制約をトラヒックの経路に反映する為に、OSPF ではこれらの PLI (Physical Layer Impairment : 物理信号劣化) を扱う為の拡張を標準化している。[15]によると、物理レイヤー損失は3つのカテゴリへと分類される。一つは線形損失と呼ばれ、ファイバー長に比例して影響を受ける。CD(Chromatic Dispersion : 色分散)やGVD(Group Velocity Dispersion : 群速度分散), PMD(Polarization Mode Dispersion : 偏波モード分散)などがこれにあたる。これらの損失はシグナル強度に依存しない、また各波長の信号に個別に影響する。エンドツーエンドの光パスに対する影響はファイバーのパラメータなどから見積もり可能である。2つ目は非線

形損失である。各チャンネル内の自己位相変調 (SPM : Self Phase Modulation) や誘導ブリルアン散乱 (SBS : Stimulated Brillouin Scattering), 誘導ラマン効果 (SRS : Stimulated Raman Scattering) などがこれにあたる。またチャンネル間のクロストークである XPM(Cross Phase Modulation : 相互位相変調) や, FWM(Four-Wave Mixing : 四波混合)[16] も同じ分類である。そして, 3つめとしてその他の分類とされる損失が, 増幅器から発生する ASE(Amplified Spontaneous Emission : 増幅自然放出)[17] などのノイズも挙げられる。従来, これらの PLI は一定値を超えた場合に信号を検出することが困難となり通信ができなくなるため, それぞれの値が上限を超えないような経路の選択は重要な課題であった。実際には2.1.4節にて述べた通り, 加算的や乗算的なメトリックを制約としたマルチメトリックでの経路計算は計算時間を必要とする。これらのメトリックの多くは距離に依存するため, 経路選択は実距離に基づいた最短経路により行い, その選択された経路の各 PLI が上限値を超えていないか確認する為に, PLI 値の情報が用いられるだろう。

近年では, デジタルコヒーレント技術により, 光の波としての性質を利用した伝送技術が一般的となり, これらの PLI 値による制約は大幅に改善された。1000km 級の長距離伝送も可能となり, また変調を行う事により通信速度も大幅に向上した。通常, 変調方式は光信号の送信機および受信機によって変更できない事が一般的であったが, Aoki ら [18] の研究では, 共用のデジタル信号プロセッサを用いる事によって, 動的な変調方式の選択を可能とした。これによって, 例えば短距離では 100Gbps, 長距離では 40Gbps といったような選択が可能となる。また, 同じ研究グループの Sone ら [19] では, DWDM(Dense WDM) において通信中の無瞬断での波長変更 (hitless と呼ばれる) による断片化の解消を実現した。この研究では, ITU-T における 100GHz 単位のチャンネル割当ではなく, より柔軟性を持ったチャンネル割当手法を必要とする。これらは, フレキシブルグリッドと呼ばれ, ITU-T[20] におい



ても現在標準化の最中である。

今後もプロセッサの発達に伴い、物理レイヤーはより柔軟性を持った技術を提供することが可能となるだろう。これらの柔軟性は選択肢の幅を広げる事を意味する。つまり、経路の選択を行う制御、および管理プレーンの役割は増えていくだろう。

### 2.1.6 伝送プレーン技術：イーサネット

MAC層の技術であるイーサネットはLANに広く提供されてきた。このイーサネットの非常に単純な転送手法は、高い市場シェアと共に、高速かつ費用対効果の高いイーサネット機器の発展を促進させてきた。1980年に10Mbpsだった当初のイーサネットは、2002年にIEEE 802.3ae[21]の標準とともに10Gbpsを達成し、現在では40Gbpsおよび100GbpsがIEEE 802.3bm ワーキンググループによって標準化されている。イーサネットは現在LANからキャリアグレード伝送ネットワーク技術としてWAN (Wide Area Network 広域ネットワーク) に使用されつつある。この用途におけるイーサネットを広域イーサネットと呼び、SONET/SDH (Synchronous Optical Network/Synchronous Digital Hierarchy) のような古くから使われてきた同期型の伝送ネットワーク技術を置き換えようとしている [22, 23]。

広域イーサネットが注目を浴びる大きなきっかけとなったのは、イーサネットの802.1Q VLAN (Virtual LAN) 拡張 [24] である。通信をブロードキャストに頼るイーサネットでは、ブロードキャストの到達範囲であるセグメント内でしか通信を行うことができない。そしてセグメントのサイズ (すなわちセグメント内の端末数) に制限を持つため、そのままでは広域ネットワークにおいて利用はできなかった。VLAN は物理的なネットワーク上に自由に複数の論理的なセグメントを構築する技術であり、VLAN ID 毎に経路を確立する事が可能となる。そして、このVLANは多

くのスイッチが標準でサポートしており、新たな開発コストを必要とせずとも広域イーサネットが構築可能であった。この VLAN を適切に設定する事によって、ネットワーク上に仮想的な接続を確立することができる。この接続を EVC (Ethernet Virtual Connection : イーサネット仮想線) と呼ぶ。EVC の利点は費用対効果の高い通信だけではなく、レイヤー 2 上のシームレスな通信にも挙げられる。特に、トンネリングなどのネットワーク機器への新たな機能の追加を必要としないため、L2VPN において非常に有用である。一方で、この VLAN を適切に設定する技術が必要となることは言うまでもない。そのため、GMPLS は広域イーサネットにおける TE 技術として大きく注目された。

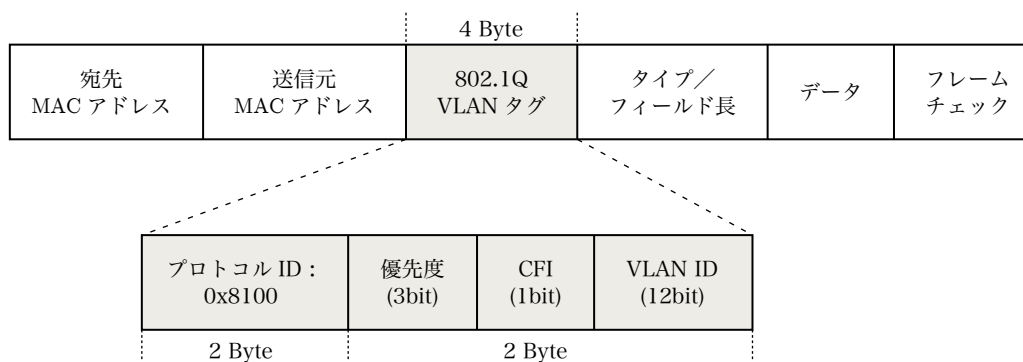


図 2.12: VLAN フレーム

## 2.2 トラフィックエンジニアリング関連研究

TE の概念が最初に導入されたのは MPLS であり [25] [26], これが GMPLS へと発展した。GMPLS では理論上, LSP によって任意の経路に任意の割合で経路を設定する事が可能であるのに対し, IP ベースの TE では任意の経路を設定する事は

不可能とされている [27]. それでも、重みの最適化だけで LSP 管理のような複雑なオーバーヘッドを必要としない点は、信頼性とスケーラビリティにおいて有利となる。その研究の多くは最適化の特性上、ほとんどがオフラインの TE であり、また通信品質の保証など複雑な機能の実現は困難である。特に、リンク障害などの場合にはネットワーク全体でトラヒックの変動が起こり、その動向は予測が困難である。[28] において指摘されるように、リンク障害によって別の経路へと避難したトラヒックが輻輳を起こす可能性は高い。

一方の GMPLS の TE はオフライン、オンラインそれぞれにおいて経路最適化が多く研究されている。図 2.1 に、経路の制御に関する IP ベースの TE と GMPLS ベースの TE についてまとめた。GMPLS ベースの TE では、経路の最適化だけでなく、遅延やパケット損失などの通信品質を 1 つのメトリックと見立てた経路選択や、DiffServ なども研究されている。更には、下位レイヤーの制約を制御するための技術なども研究された。これらについても紹介していく。

表 2.1: IP ベースの TE および GMPLS ベースの TE

	GMPLS ベースの TE	IP ベースの TE
経路決定メカニズム	フローベースに明示的	送信先ごとのホップ毎
マルチパス伝送	任意に可能	等分分配 (ECMP) のみ
経路選択の柔軟性	高い (任意に経路を設定)	低い (最短経路のみ)

### 2.2.1 IP ベースのトラフィックエンジニアリング

IP ベースの TE は、実際には MPLS ベースの TE の代替案として、MPLS の TE よりも後に提案された [29]。コネクションを持たない IP において、重みの最適化により宛先ベースのルーティングだけで TE を実現する。

先に述べた ECMP では、同じ宛先のトラフィックは等分に分割され、複数の経路を通過する。実際にはトラフィックの分割は実際にはパケット単位ではなく、フロー単位で行われる。なぜなら、パケットの順序逆転を避ける為に、同じ送信元からのパケットが複数の経路に分割される事を避ける為である。これらのマルチパスアプローチは、Netscope TE ツール [30] に最初によって提案され分析された。

リンク重みの最適化を用いた ECMP による負荷分散は Forts および Thorup によって提案された [29, 31, 32]。従来の、帯域の逆数となる重み付けに比べ、50 から 110 パーセントのネットワーク容量改善が見られた。この手法では、ECMP を実現するリンク重みの最適化問題を ILP(Integer Linear Programming : 整数線形計画問題) へと帰着させ、この問題を ILP ソルバーと呼ばれるソフトウェアによって解いた。Forts および Thorup によって証明されるように、最適なリンク重みの導出は NP 困難であると知られており、高性能なソルバーを用いたとしても非常に長い時間を要した。そこで、彼らは局所探索法によるアプローチも用いている。[33] しかしながら、この方式では解空間の限られた部分空間ごとにしか探索が行われないため、発生した乱数による解のばらつきが大きい。一方、Ericsson [35] は GA(Genetic Algorithm : 遺伝的アルゴリズム) によるアプローチを提案した。GA によるアプローチでは、ILP ソルバーを用いるよりも早い段階での収束が得られ、また解空間全体を考慮した探索が可能となる。そのため、非常に良い解が短時間で求められる。

しかしながらそれでもこの複雑な計算は、ノードサイズの拡大とともに計算時間が莫大となる。一方の計算を実行する CPU に注目すると、その動作速度の向上は

期待できない。動作周波数の向上による排熱問題から、CPU ベンダーはクロック速度の高速化からマルチコア化に方針を転換している。すなわち、1つのタスクはそのままでは高速化されず、並列に処理することが必要となることを意味する。本論文ではこのリンク重みの最適化計算を、昨今のマルチコア CPU において高速化する為の手法を提案した。これらは第3章にて詳述される。

一方、最適な TE のためにはトラヒック分割が必須であるという考え方から、ECMP ではなく不均等なトラヒック分割を設定するアプローチも研究された。Sridharan らの研究 [36] では、既存のプロトコルを一切変えずに IP プレフィックスを用いてより効率的な TE を実現した。この方式では、ネットワークプレフィックスにより集約したトラヒックのルーティングと個々の IP アドレスによって経路を使い分ける事によって効率的な TE を実現する。MPLS がフローの送信ノードと受信ノードに対応するラベルを使用する代わりに、この方式では送信先アドレスのネットワークプレフィックスを受信ノードアドレス、そしてホスト番号をラベルと使っている事にすぎず、IP アドレス空間が十分に広い場合は良いが、IP の枯渇を早める事となることが予想される。

J. Wang らの研究 [37] では、Edge ベースのフロー分割を用いた手法を提案している。この手法では、予め物理ネットワークをいくつかの論理ネットワークに分離する。ECMP を利用しない代わりに、ネットワークのエッジノードに到着するフローを不均一に分割し、いずれか論理ネットワーク上で送信される。各フローがどの論理ネットワークにて送信されているかは、IP ヘッダ内の ToS (Type of Service) 値を利用して識別される。MPLS の不均等なフロー分割を IP ネットワーク上で再現したこの手法において最も大きな課題は、それぞれの論理ネットワークに対する重み付けの計算である。この研究ではシミュレーションにより、だいたい2から4つの論理ネットワークにより、十分な TE が達成されることを示している。

## 2.2.2 GMPLS ベースのトラヒックエンジニアリング

### 2.2.2.1 GMPLS ベースのオフライントラヒックエンジニアリング

GMPLS におけるネットワークでは明示的な LSP によって通信が行われるため、フロー毎の経路を指定した TE が可能である。このフローとは始点ノードと終点ノードが同一であるトラヒックに対して複数指定する事が可能である。即ち、トラヒック分割が任意のノードにて任意の割合で行うことができる。この経路の最適化はフロー分割を含むマルチコモディティフロー問題として LP (Linear Programming : 線形計画問題) に定式化される [38]。これらの式は、LP ソルバーと呼ばれるソフトウェアによって多項式時間内に解く事ができる。つまり、GMPLS における経路の最適化計算は NP 困難ではない。

しかし、このような最適な経路選択による TE は別の問題を持つ。任意のフロー分割により、ネットワークサイズが大きい場合に大量の LSP 確立を要求する。実際ネットワークでは LSP を制御するための実装コストが存在するため、確立可能な LSP 数は限られており、このアプローチは実用的ではない。即ち、GMPLS におけるオフライン TE は、LSP の制御および管理問題に帰着すると言える。

ここで、LSP の管理を簡単にするために、LSP が1つの送信ノードから一つの宛先ノードへと確立されるという P2P (Point-to-Point) モデルから少し離れる。複数のノードから一つの宛先への MP2P (Multipoint-to-Point) のツリー型の LSP を複数確立し、送信ノードではこれらの LSP を選択して送信することにより TE を行う方式が提案された [39, 40, 41]。これらのモデルでは、LSP の確立および管理におけるスケーラビリティの問題が緩和される。LSP の数を減らすという目的から [39] は複数の MP2P を用いる手法を提案した。厳密には、MP2P のパスを確立する手法と、MP2P のモデルによってトラヒックフローを割り当てる手法の両方である。この手法では、まず P2P のパスがアルゴリズムによって選択され、これらをバイナリ線形

計画問題による選択問題として解く。この方法により、負荷分散を行ったまま LSP 管理のオーバヘッドを減少させる。また、[40] では QoS 保証のための TE が MP2P によって実現された。[41] ではラベル空間に関するスケーラビリティについて検証が行われた。これらの提案は P2P をマージするという考え方に基づいており、課題は MP2P への割当だけである。そのため、経路やリソース割当の最適化は主たる問題ではない。

### 2.2.2.2 GMPLS ベースのオンライントラヒックエンジニアリング

2.1.3 節にて紹介したように、予測したトラヒック行列に対して最適な経路を計算し割当て、それを適用するというモデルに基づいた TE がオフライン TE である。しかし、実際のネットワークにおける通信の要求の全てがこのモデル当てはまるとは限らない。特に、コアネットワークでは VPN (Virtual Private Network) のように特定ノード間での固定的な通信帯域を要求するようなサービス要求も少なくない。その場合、トラヒック行列が予測できない代わりに、フロー単位でのトラヒック要求が行われる。これを動的に一つずつネットワークへとフローとして割り当てるような TE が求められる。

帯域を考慮したフロー単位の経路選択では CBR (Constraint-Based Routing : 制約ベースルーティング)[45] が一般的である。この方式では、2.1.4 節にて紹介した CSPF というアルゴリズムに基づく。帯域が不十分など、条件に合わないリンクを先にグラフ上から消去してから SPR (Shortest Path Routing : 最短経路探索) を行い、経路を決定する。似たような手法として、WSP (Widest Shortest Path : 広帯域最短経路探索)[46] や SWP (Shortest Widest Path 広帯域最短経路探索)[47] も提案されている。両者はトラヒックを割り当てた後の有効帯域がボトルネックとならぬようルーティングを行うものである。つまり、未来のトラヒックのための帯域を確保しておく

ことによるアプローチによって、ボトルネックを解消しようという戦略である。これらは、それぞれのフローに対する独立したルーティングの手法であり、それぞれのトラヒック要求は将来のトラヒック要求に関係なく確立される。

これらの手法は到着順割当 (First-come-First-serve) と呼ばれ、より早く要求されたフローに優位性がある。ネットワークの最適化のためには LSP の張り替えが必要となるが、新たな LSP が確立される度にいくつもの LSP の張り替えを伴うのは現実的ではない。ここで、MATE [43, 44] では珍しいアプローチをとった。この手法では図 2.13 のように予め静的に作成された LSP へのトラヒックフローの分割割合を動的に変化させることによって TE を実現する。この手法では、リソースの最適化がネットワーク上で実現される為、経路最適化は直接関与しない。

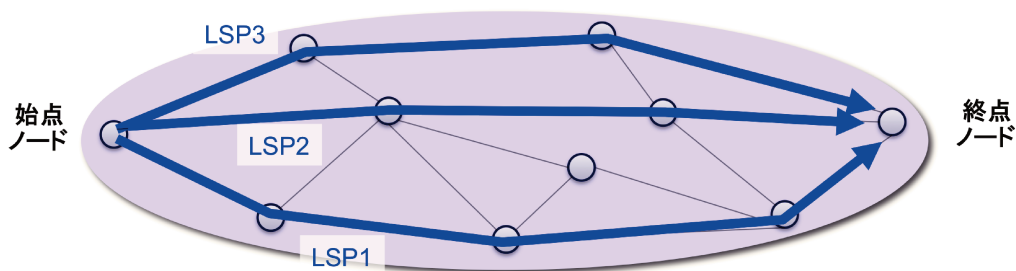


図 2.13: MATE: MPLS Adaptive Traffic Engineering の概念

しかしながら、MATE では選択される経路によってトラヒックの遅延値が異なるため、順序逆転などが懸念される。そこで、DORA [48] では将来の以降のトラヒック要求の可能性を考慮した手法を提案した。この手法では 2 段階のアルゴリズムを用い、以降に到着するトラヒック要求を収容する可能性を最大化しながら、トラヒック要求を割り当てる。最初のステージでは、各リンクに対し始点ノード/終点ノードペア毎にパスポテンシャル値と呼ばれる値を設定する。この値は、全ての始点ノード



ド/終点ノードペアの K-shortest ルーティングによるディスジョイントパス上でそのリンクが何回使用されているかを表した値に基づいており、そのリンクが使用される可能性を示す。そして、トラヒック要求が到着する度に、この値と、各リンクの残余帯域の積として重み付けたグラフで CSPF アルゴリズムを実行し、経路を割り当てる。

MPLS ベースのオンライン TE において重要な課題として、いくつかの研究では LSP の干渉に注目した [49, 50, 51, 52]。LSP 干渉とは、特定のノードペア間の経路として重要なリンクにおける LSP の競合である。ここで重要なリンクとは、そのリンクが使用できない場合に遠回りを強いられるリンクである。もしもパスを確立しようとしている経路上のあるリンクが、他のノードペア間に確立する LSP にとって重要なリンクである場合、そのリンクの使用を避け、別の経路を使用した方が、競合を避けるには良い。この考えに基づいて、[49, 50] では、MIRA (Minimum Interference Routing Algorithm : 最小干渉ルーティングアルゴリズム) を提案している。まず、LSP 干渉の起こりやすいクリティカルリンクを決定する為に、各ノード間での最大フローを計算し、この値に基づき、クリティカルリンクセットを求める。このセットから各リンクの重み付けを行い、CSPF アルゴリズムを適用する事で、各トラヒック要求を処理する。これらの研究では、Routing and Traffic Engineering Server (RATES) と呼ばれるソフトウェアを実装したものが報告されている [51]。

### 2.2.2.3 GMPLS ベースの DiffServ 環境におけるトラヒックエンジニアリング

ここまで、経路の最適化による輻輳の解決だけに焦点を当てた TE について紹介してきた。一方で、MPLS によって QoS 要求を満たすネットワークに関する TE についてここで紹介する。主に、DiffServ (Differentiated Services : 差別化サービス) に基づいた GMPLS ベースの TE を実現する研究である。DiffServ 認識型のオフライ

ン TE およびその類いの TE はいくつか提案されている。[53, 54, 55] トラヒックの最適化という概念から、TE の目的が QoS 要求を満たしながら、ネットワーク全体のコストを最小化することである。そこでコスト関数を各 QoS クラスに対してそれぞれ用意し、ネットワーク全体での TE をコンベックス空間における非線形計画問題として表した。[54] では GDP (Gradient Projection) アルゴリズムによってこれを解く事によって DiffServ 認識型の TE を実現した。ここで QoS は、遅延およびパケットロスであり、両者はホップ数の制約として表された。[55] では DTE (Differentiated TE : 差別化 TE) が提案された。この DTE では全ての経路最適化が部分問題へと細分化され、解空間がコンベックスとならない部分問題は焼き鈍し法にて、それ以外は GDA によって解決された。

一方、DiffServ 環境におけるオンライン TE としては、優先度の高いサービスにおける LSP 要求の為に、優先度の低い LSP の張り替えを行う TEAM [56] が提案された。頻繁な張り替えによるネットワークの不安定化を避ける為に、提案される張り替えのアルゴリズムは、低い優先度の張り替え、少ない数の張り替え、少ない帯域の張り替えを率先して行う。LSP のルーティングには、SPeCRA (Stochastic Performance Comparison Routing Algorithm) [57] が TEAM において使用される。SPeCRA ではマルコフ連鎖を用い、状況に応じてルーティングアルゴリズムを変更し、最適なアルゴリズムでルーティングを行う。SPeCRA の著者は Diffserv ベースの LSP 張り替えポリシーを V-PREPT として提案した [58]。V-PREPT では TEAM 同様、LSP の優先度、張り替えする LSP 数、帯域の 3 つの基準が調整される。ネットワークの堅牢性 (survivability) に注目したオンライン TE は、K. Walkowiak によって提案された [59]。MIRA と同様に、このスキームではリンクへの独自の重みに基づいて最短経路探索を行い LSP を動的に確立する。この動的なリンク重みは LFL (Lost Flow in Link) 関数によって与えられる。

#### 2.2.2.4 GMPLS ベースのマルチメトリック環境におけるトラヒックエンジニアリング

GMPLS では DiffServ 環境のような複雑な TE も LSP によって制御可能である。しかしその一方で、「優先度」や「張り替え頻度」などの新たな値がネットワーク上で交換されなければならない。これらの多様なメトリック（評価）値を用いた TE を、マルチメトリック TE と呼ぶ。先に述べた通り、GMPLS は拡張によりマルチメトリックがサポートされる。

マルチメトリック TE として最も一般的な技術が、後に詳述される物理レイヤーの評価である。その中で最も標準化に近いものが PLI 評価である。[60] この拡張では、2.1.5 節において述べたような PMD や CD, ASE などの光信号の情報をメトリック情報として扱う事により、光信号の信号劣化を考慮した経路選択が可能となる。PLI のための拡張に関する標準化は現在、IETF により行われている。[61] [62] また [63] では、リンク損失率をメトリックとして考慮した信頼性の高い LSP 確立のための拡張に向けた研究がされている。GMPLS が対応すべきメトリック情報は、ネットワークに関するメトリックだけではない。近年のいくつかの研究では、電力消費量を考慮したエコロジーネットワークとして、電力量をメトリックとする拡張が検討されている [64] [65]。ユビキタスネットワーク環境では、様々なサービスやアプリケーションに関する情報が LSP 確立において考慮される [66]。また、無線の分野ではモバイルバックホールネットワークにおける一つの基地局あたりの端末数を考慮した LSP 確立、CDN (Contents Delivery Network : コンテンツ配信ネットワーク) ではコンテンツサーバーあたりのユーザー数、その他の要素など様々なサービスに関係するメトリックが GMPLS によって対応される必要がある。このように、マルチメトリックにより様々な値が OSPF-TE によって交換が必要になった場合、この値を交換する制御システム自身のスケーラビリティに問題が生じる。これについて

は第4章にて扱われる。

## 2.2.3 下位レイヤーに関連したトラヒックエンジニアリング技術

### 2.2.3.1 広域イーサネット

既に紹介したように、イーサネットはLANにおける市場とともに、高速かつ費用対効果の高い通信技術を確立した。現在ではキャリアネットワークにおいてSONET/SDHの技術を置き換えようとしている。しかしながら、キャリアグレードの広域ネットワーク技術として利用するには、イーサネットはいくつかの機能に欠けている。

最初に必要となるのは制御プレーンであった。イーサネットはその転送方式からスター型のトポロジーしか構成できず、閉回路を作った場合にはネットワークは崩壊してしまう。STP[67]はイーサネット上の閉回路を認識し、いずれかのリンクの切断する事によってスター型を保持していた。切断された方のリンクは、リンク障害時用のバックアップ回路として利用する事ができる。しかしながら問題はその収束時間であった。リンク障害が発生した際に、切り替えまでに1分近い収束時間が発生する。これを改善したのがRSTP[68]である。RSTPでは代替ポートとバックアップポートを追加する事によって、収束時間を数秒とした。これらのプロトコルは、リンクの冗長性としては有効であったが、経路の切り替えには使用できなかった。[69]はコネクションの制御のためにGMPLSを用いて制御する方法を提案した。GELS(GMPLS-controlled Ethernet Label Switching)[70]はコンピュータシミュレーションにより、GMPLS制御のイーサネットが、従来のRSTPベースのツリー構築に比べて45パーセントも多く帯域を利用できる事を示した。また[71]ではGELSに必要とされるGMPLS拡張について言及している。

GMPLSにより経路の制御が自由になった一方で、次の問題はID空間の管理であった。VLANによるネットワーク内での経路の制御は、ネットワークの始点ノードにてVLAN IDを付与する事が前提である。即ち、既にVLAN IDが付与されたフレームを処理する事が不可能であった。この問題を解決する為に、VLAN IDのスタック技術 [72] が提案された。この技術はプロバイダブリッジと呼ばれ、キャリアネットワーク内において別のVLAN ID空間とする。さらに、より深いIDのスタックと、アドレス空間の分離はプロバイダバックボーンブリッジ [73] として提案された。しかしながら、これらは複数のドメイン間におけるIDの競合の問題については考慮していない。そこで、本研究の第5章において、ドメイン間におけるVLAN IDのスイッチを実現するための制御技術について提案した。

一方で、広域イーサネットを用いたサービスは、メトロイーサネットフォーラム [74] において議論されている。イーサネットにおける多点間接続技術はそのサービスの一つとして要求されている。本論文の第6章では、多点間接続技術するためのGMPLSシグナリングについて取り扱う。

## 2.3 本論文の位置づけ

ここまで述べたTEを実現するための関連技術と、TEに対するこれまでの研究を次の図2.14にまとめた。まずTEの戦略によって、単に経路の交換によって通信の効率性をあげるものと、下位レイヤーとの互換性を解消して通信の効率性を上げるものに大別できる。トラヒックの経路を交換する場合、そのネットワークの種類によって、IPベースとMPLSベースに分けられ、MPLSにおいてはオフラインで行うもの、オンラインで行うもの、そして特定の条件においてのみTEが実現されるものに区別した。一方の下位レイヤーについては、イーサネットとWDMネット

ワークにおいて、その特徴を考慮した方式により通信の効率性を上げるための研究がある。以降の章において述べる本論文における研究貢献は、同じく図2.14のようになる。各章における目的、課題、提案および成果は、表2.2の通りである。

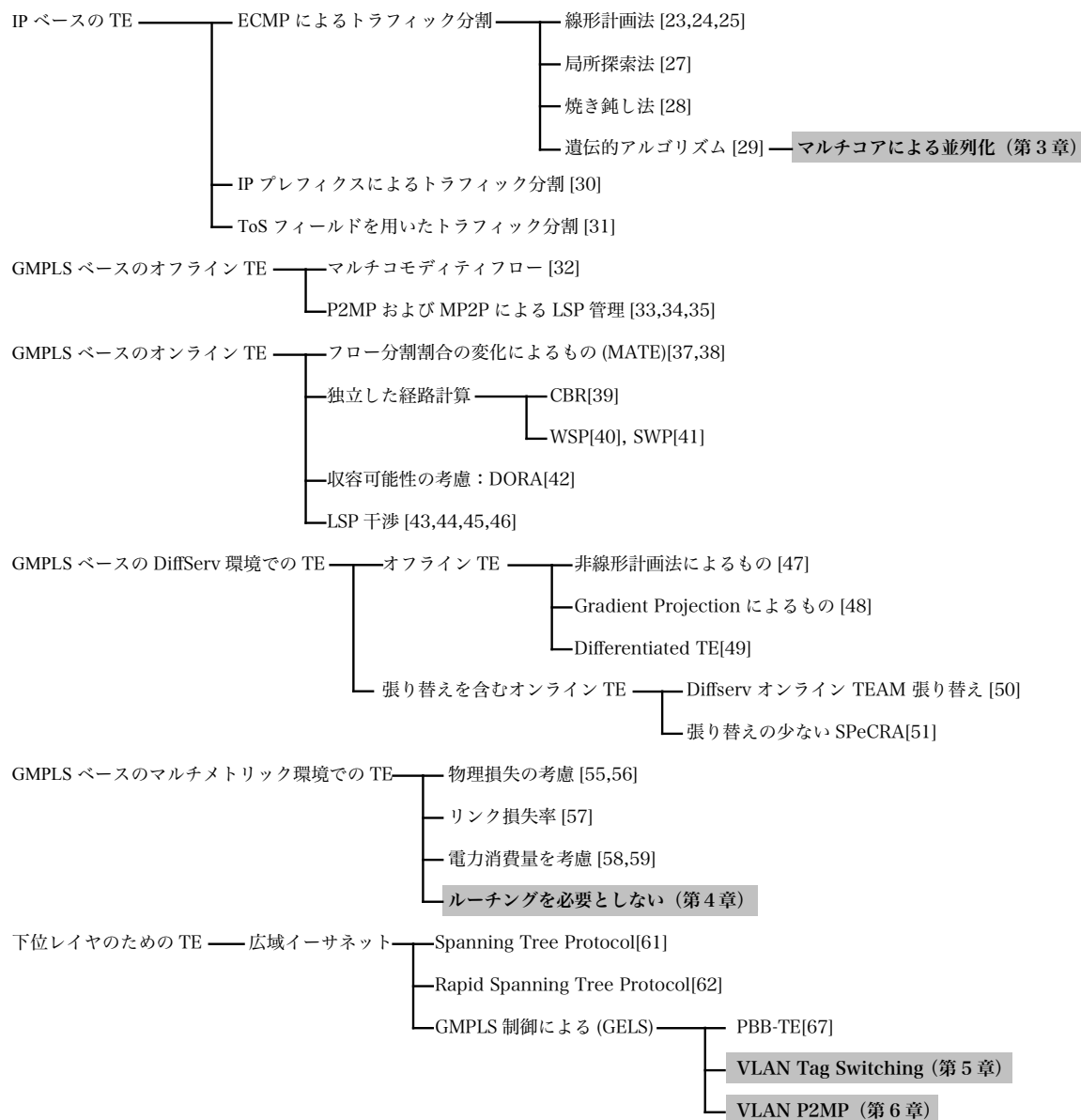


図 2.14: 各章の研究の位置づけと要約

表 2.2: 各章における研究の課題と提案

第3章	目的	IP ベース効率的な TE におけるネットワークのパラメータの最適化による輻輳の解決を図る.
	課題	最適化計算は計算量が大きく短時間で行う為の計算速度の高速化が必要となる. マルチコア CPU における並列処理による高速化のための効率的な並列アルゴリズムが要求される.
	提案	データのコヒーレンスを考慮し, 計算の効率性を向上したデータ配置およびアルゴリズムを提案した.
	成果	16 コアのマルチコア CPU において, 単に従来方式では約 7 倍程度の高速化に対し, 提案アルゴリズムでは 13 倍まで高速化が可能となった.
第4章	目的	GMPLS ベースの TE において, 距離だけでなくパケット損失率や消費エネルギーなど様々なネットワーク要求を考慮するために, メトリックと呼ばれる値の多様化をサポートする.
	課題	多様化したメトリックの更新メッセージの爆発的な増加により, ネットワーク上の通信帯域や処理リソースが奪われるためこれを回避しなければならない.
	提案	ルーティングプロトコルを用いずメッセージの配布を行わずに, 経路導出とパス確立を同時に行うシグナリングベースの FB-RSVP-TE プロトコルを提案した.



	成果	メトリック更新に対するメッセージ数な増加を防ぎ、メトリック数に多様化を許容するネットワークを実現した。
第5章	目的	広域イーサネットにおけるマルチドメイン環境のTEを実現する。
	課題	本来はLANでの運用を前提としたイーサネットを広域に拡大した場合、ID空間の制限によりネットワーク全体でコネクション数に制約を生じる。
	提案	シグナリング機能を拡張し、ドメイン境界ノードにおいてタグをスワップさせることにより、ドメイン毎のID空間を独立化を行った
	成果	各ドメインでのID空間による制約は残るものの、ネットワーク全体ではその制約から開放された。
第6章	目的	広域イーサネットにおける効率的なマルチキャストサービスを実現する
	課題	広域イーサネットにおける多点間接続を確立するための、受信側を接続する経路の指定を含んだシグナリングの実現。
	提案	受信側端末のVLAN IDを指定する新たなオブジェクトをRSVP-TEメッセージを新たに定義した。
	成果	プロトタイプネットワーク上でイーサネットによる多点間接続技術が実現された

## 参考文献

- [1] G. Malkin, Editor, “RIP Version 2,” IETF Internet Standard, RFC 2453, November 1998.
- [2] Moy, J., “OSPF Version 2”, IETF Internet Standard, RFC 2328, April 1998.
- [3] G. Iannaccone, et al., “Analysis of Link Failures in an IP Backbone,” Proceedings of ACM IMW, Page 237-42, November 2002.
- [4] Ghazala, A.A., et al., “A Survey for Open Shortest Path First Weight Setting (OSPFWS) Problem, ” Proceedings of ISA 2008, Page 111-116, April 2008.
- [5] K. Kompella, et al., “OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS),” IETF Proposed Standard, RFC 4203, October 2005.
- [6] D. Awduche, et al., “RSVP-TE: Extensions to RSVP for LSP Tunnels,” IETF Proposed Standard, RFC 3209, December 2001.
- [7] D. Awduche, et al., “Traffic Engineering (TE) Extensions to OSPF Version 2,” IETF Proposed Standard, RFC 3630, September 2003.

- [8] J. Lang, et al., "Generalized Multi-Protocol Label Switching (GMPLS) Recovery Functional Specification," IETF Proposed Standard, RFC 4426, March 2006.
- [9] Marzo Jose L., et al., "QoS online routing and MPLS multilevel protection: a survey," IEEE Communications Magazine, Volume 41, No. 10, Page 126-132, October 2003.
- [10] A. Asgari, et al., "Scalable Monitoring Support for Resource Management and Service Assurance," IEEE Network, Volume 18, No. 6, Page 6-18, December 2004.
- [11] P. Trimintzios, et al., "A Management and Control Architecture for Providing IP Differentiated Services in MPLS-Based Networks," IEEE Communication Magazine, Volume 39, No. 5, Page 80-88, May 2001.
- [12] Z. Wang, et al., "Quality of Service Routing for Supporting Multimedia Applications," IEEE Journal on Selected Areas in Communications, Volume 14, No. 7, Page 1228-34, September 1996.
- [13] A. Farrel, et al., "A Path Computation Element (PCE)-Based Architecture," IETF Informational Documents, RFC 4655, August 2006.
- [14] Rahbar, Akbar Ghaffarpour, "Review of Dynamic Impairment-Aware Routing and Wavelength Assignment Techniques in All-Optical Wavelength-Routed Networks," IEEE Communications Surveys and Tutorials, Volume 14, Issue 4, March 2009.

- [15] G. Agrawal, "Fiber-optic communication systems, 3rd edition," Volume 1, Wiley New York, 2002.
- [16] M. Lima, A. Cesar, and A. Araujo, "Optical network optimization with transmission impairments based on genetic algorithm," in IEEE International Microwave and Optoelectronics Conference, Brazil, Page361-365, September 2003.
- [17] E. Desurvire, D. Bayart, B. Desthieux and S. Bigo, "Erbium-doped fiber amplifiers: Device and System Developments," Wiley-Interscience, July 2002.
- [18] Aoki Yasuhiko, et al., "Dynamic and Flexible Photonic Node Architecture with Shared Universal Transceivers Supporting Hitless Defragmentation," Optical Society of America, Proceedings of International Conference on European Conference and Exhibition on Optical Communication(ECOC), No. We-3-D-2, September 2012.
- [19] Sone Kyosuke, et al., "First demonstration of hitless spectrum defragmentation using real-time coherent receivers in flexible grid optical networks," Optical Society of America, Proceedings of International Conference on European Conference and Exhibition on Optical Communication(ECOC), No. Th-3-D-1, September 2012.
- [20] ITU-T G.694.1 Spectral grids for WDM applications: DWDM frequency grid, June 2012.
- [21] IEEE Computer Society, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks- Specific Requirements Part 3: Carrier Sense Mul-

- multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment: Media Access Control (MAC) Parameters, Physical Layers, and Management Parameters for 10 Gb/S Operation,” IEEE Standards 802.1ae, August 2002.
- [22] Fouli, K. ; Maier, M., “The road to carrier-grade Ethernet,” IEEE Communications Magazine, Volume 47, Issue 3. Page S30-S38, March 2009.
- [23] Reid, A., et al., “Carrier ethernet,” IEEE Communications Magazine, Volume 46, Issue 9, Page 96-103, September 2008.
- [24] IEEE Computer Society, “IEEE Standards for Local and Metropolitan Area Networks Virtual Bridged Local Area Networks,” IEEE Standard 802.1Q, May 2006.
- [25] D. Awduche, et al, “Requirements on Traffic Engineering over MPLS,” IETF Informational Document, RFC 2702, June 1999.
- [26] D. Awduche, et al., “MPLS and Traffic Engineering in IP Networks,” IEEE Communication Magazine, Volume 37, No. 12, Page 42-47, December 1999.
- [27] Ning Wang, Kin Ho, Pavlou G., Howarth M., “An overview of routing optimization for internet traffic engineering,” Communications Surveys and Tutorials, IEEE, Volume 10, Issue 1, Page 36-56, April 2008.
- [28] B. Fortz, et al., “Robust Optimization of OSPF/IS-IS Weights,” Proceedings of INOC 2003, Page 225-30, October 2003.
- [29] D. Katz, et al., “Internet Traffic Engineering by Optimizing OSPF Weights,” Proceedings of IEEE INFOCOM, Page 519-28, March 2000.

- [30] A. Feldmann, et al., “NetScope: Traffic Engineering for IP Networks,” IEEE Network, Volume 14, No. 2, Page 11-19, Mar./Apr. July 2000.
- [31] B. Fortz, et al., “Optimizing OSPF/IS-IS Weights in a Changing World,” IEEE Journal on Selected Areas in Communications, Volume 20, No. 4, Page 756-67, May 2000.
- [32] B. Fortz, et al., “Traffic Engineering with Traditional IP Routing Protocols,” IEEE Communicaiton Magazine, Volume 40, No. 10, Page 118-24, October 2002.
- [33] B. Fortz, et al., “Increasing internet capacity using local search. ” Journal of Computational Optimization and Applications, Volume 29, Issue 1, pp 13-48, October 2000.
- [34] Mohammed H. Sqalli, Sadiq M. Sait, and Mohammed Aijaz Mohiuddin “An Enhanced Estimator to Multi-objective OSPF Weight Setting Problem”, 2006
- [35] M. Ericsson, et al., “A Genetic Algorithm for the Weight Setting Problem in OSPF Routing,” J. Combinatorial Optimization, Volume 6, No. 3, Page 299-333, September 2002.
- [36] A. Sridharan, et al., “Achieving Near-Optimal Traffic Engineering Solutions for Current OSPF/IS-IS Networks,” IEEE/ACM Transactions on Networking, Volume 13, No. 2, Page 234-47, April 2005.
- [37] J. Wang, et al., “Edge Based Traffic Engineering for OSPF Networks,” Computer Networks, Volume 48, No. 4, Page 605-25, July 2005.

- [38] D. Mitra and K. G. Ramakrishnan, "A Case Study of Multiservice, Multi-priority Traffic Engineering Design for Data Networks," Proceedings of IEEE GLOBECOM 1999, Page 1077-83, December 1999.
- [39] H. Saito, et al., "Traffic Engineering Using Multiple Multipoint-to-point LSPs," Proceedings of IEEE INFOCOM, Page 894-901, March 2000.
- [40] G. Urvoy-Keller, et al., "Traffic Engineering in a Multipoint-to-Point Network," IEEE Journal on Selected Areas in Communications, Volume 20, No. 4, Page 834-49, May 2002.
- [41] S. Bhatnagar, et al., "Creating Multipoint-to-Point LSPs for Traffic Engineering," IEEE Communications Magazine, Volume 43, No. 1, Page 95-100, January 2005.
- [42] P. Trimintzios, et al., "Engineering the Multi-Service Internet: MPLS and IP-Based Techniques," Proceedings of IEEE ICT, Page 129-34, June 2001.
- [43] A. Elwalid, et al., "MATE: MPLS Adaptive Traffic Engineering," Proceedings of IEEE INFOCOM, Page 1300-09, April 2001.
- [44] S. Kandula, et al., "Walking the Tightrope: Responsive Yet Stable Traffic Engineering," ACM SIGCOMM Computer Communications Review, Volume 35, No. 4, Page 253-64, October 2005.
- [45] O. Younis, et al., "Constraint-Based Routing in the Internet: Basic Principles and Recent Research," IEEE Communications Surveys and Tutorials, 3rd qtr., Page 2-13, May 2003.

- [46] Z. Wang, et al., "Quality of Service Routing for Supporting Multimedia Applications," *IEEE Journal on Selected Areas in Communications*, Volume 14, No. 7, Page 1228-34, September 1996.
- [47] R. Guerin, et al., "QoS Routing Mechanisms and OSPF Extensions," *Proceedings of IEEE GLOBECOM 1997*, Page 1903-08, November 1997.
- [48] R. Boutaba, et al., "DORA: Efficient Routing for MPLS Traffic Engineering," *J. Network and Systems Mgmt.*, Volume 10, No. 3, Page 309-25, September 2002.
- [49] K. Kar, et al., "Minimum Interference Routing of Bandwidth Guaranteed Tunnels with MPLS Traffic Engineering Applications," *IEEE Journal on Selected Areas in Communications*, Volume 18, No. 12, Page 2566-79, December 2000.
- [50] K. Kodialam, et al., "Minimum Interference Routing of Applications to MPLS Traffic Engineering," *Proceedings of IEEE INFOCOM*, Page 884-93, 2000.
- [51] P. Aukia, et al., "RATES: A Server for MPLS Traffic Engineering," *IEEE Network*, Volume 14, No. 2, Page 34-41, Mar./Apr. 2000.
- [52] F. Blanchy, et al., "A Preemption-Aware On-line Routing Algorithm for MPLS Networks," *Telecommunication Systems*, Volume 24, No. 2-4, Page 187-206, October 2003.
- [53] F. Le Faucheur, et al., "Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering," *IETF Informational Documents*, RFC 3564, July 2003.



- 
- [54] P. Trimintzios, et al., “Quality of Service Provisioning through Traffic Engineering with Applicability to IP Based Production Networks,” *Journal of Computer Communications*, Volume 26, No. 8, Page 845-60, May 2003.
- [55] V. Tabatabaee, et al., “Differentiated Traffic Engineering for QoS Provisioning,” *Proceedings of IEEE INFOCOM*, Page 2349-59, March 2005.
- [56] C. Scoglio, et al., “TEAM: A Traffic Engineering Automated Manager for Diff-Serv Based MPLS Networks,” *IEEE Communications Magazine*, Volume 42, No. 10, Page 134-45, October 2004.
- [57] J. C. de Oliveira, et al., “SPeCRA: A Stochastic Performance Comparison Routing Algorithm for LSP setup in MPLS Networks,” *Proceedings of IEEE GLOBECOM*, Page 2190-94, November 2002.
- [58] J. C. de Oliveira, et al., “New Preemption Policies for DiffServ Aware Traffic Engineering to Minimize Rerouting in MPLS Networks,” *IEEE/ACM Transactions on Networking*, Volume 12, No. 4, Page 733-45, August 2004.
- [59] K. Walkowiak, “Survivable Online Routing for MPLS Traffic Engineering,” *Proceedings of QofIS*, Page 288-97, September 2004.
- [60] Siamak Azodolmolky, et al., “A Survey on Physical Layer Impairments Aware Routing and Wavelength Assignment Algorithms in Optical Networks,” *Computer Networks and Isdn Systems*, Volume 53, No. 7, Page 926-944, May 2009.
- [61] Y. Lee, et al., “A Framework for the Control of Wavelength Switched Optical Networks (WSO) with Impairments,” *IETF Internet draft*, draft-ietf-ccamp-wson-impairments-10, January 2012.

- [62] Y. Lee, et al., “Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks,” IETF Internet draft, draft-ietf-ccamp-rwa-info-19, November 2013.
- [63] Luis Velasco, et al., “Failure Aware Diverse Routing: A Novel Algorithm to Improve Availability in ASON/GMPLS Networks,” Transparent Optical Networks 2006 International Conference, Page 195-198, June 2006.
- [64] Haruka Yonezu, et al., “QoS Aware Energy Optimal Network Topology Design and Dynamic Link Power Management,” European Conference and Exhibition on Optical Communication 2010, Tu.3.D.4, September 2010.
- [65] Antonio Cianfrani, et al., “An Energy Saving Routing Algorithm for a Green OSPF Protocol,” INFOCOM IEEE Conference on Computer Communications Workshops 2010, May 2010.
- [66] Daisuke Ishii, et al., “A Novel IP Routing/Signaling Based Service Provisioning Concept for Ubiquitous Grid Networking Environment,” IEEE Globecom 2010 Workshop on Ubiquitous Computing and Networks, No. 4, Page 1808-1812, December 2010.
- [67] IEEE Computer Society, “802.1D Standard for local and metropolitan area networks - Media Access Control (MAC) Bridges,” IEEE Standard 802.1D, June 2004.
- [68] IEEE Computer Society, “Rapid Reconfiguration of Spanning Tree,” IEEE Standard 802.1w, June 2004.

- [69] D.Papadimitriou, et.al, “Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE)”, IETF Proposed Standard, RFC 6003, April 2008.
- [70] Ilyas, S.M., et.al, “A simulation study of GELS for Ethernet over WAN”, Proceedings of GLOBECOM 2007, Page 2617-2622, November 2007.
- [71] D. Fedyk, et.al, “Generalized Multiprotocol Label Switching (GMPLS) Ethernet Label Switching Architecture and Framework”, IETF Informational Documents, RFC 5828, April 2008.
- [72] IEEE Computer Society, “Standards for Provider Bridges, ” IEEE Standard 802.1ad, August 2005.
- [73] IEEE Computer Society, “802.1ah - Provider Backbone Bridges (Draft 3.4),” IEEE Standard 802.1ah, April 2008.
- [74] Metro Ethernet Forum, “Metro Ethernet Networks - A Technical Overview, ” March 2004.

---

## Chapter 3

---

### IP ベースの

### トラフィックエンジニアリングにおける

### 最適化計算の規模拡張性

#### 3.1 概要

第2章では、インターネットに用いられるネットワーク技術と、TEの必要性そして関連研究について紹介した。制御プレーンの通信方式はコネクションを持たない通信であるIPと、コネクションに基づく通信のGMPLSに大別される。本章では、前者であるIPによる通信においてTEを実現する、IPベースのTEに注目する。IPベースのTEでは、OSPFメトリックに重み付けを行う事によりTEを実現するが、最適な重み付けの計算はNP困難であり、ネットワーク規模拡張性に欠ける。そこで本章では、ネットワークサイズが大きなネットワークにおいても、短時間に計算を行う手法について研究を行った。伝統的な手法である線形計画法による解法と対比して、Ericssonらの提案する遺伝的アルゴリズムを用いたヒューリスティックな解法に注目し、このアルゴリズムの計算をさらに短時間で解く為に、並列プロセッサ上で効率的な並列を実現する手法を提案した。提案手法では、データの依存関係を解消することにより効率的に処理を行うアルゴリズムを提案した。こ

れにより、16 コアの CPU において、13 倍近く of 速度向上を得た。この結果は、今後、大規模な並列処理が主流となる時代に対する大きな展望を示すものである。

## 3.2 OSPF メトリック重み付け問題

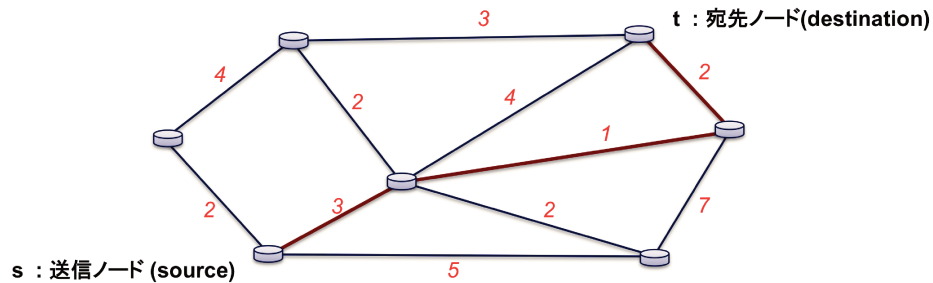


図 3.1: メトリックと最短経路

OSPF を利用するネットワークでは、メトリックに重み付けを行う事によってネットワーク内のトラヒックの経路を制御する。メトリックとは、各リンクに割り当てられる 16bit の値であり 1 から 65535 のいずれかの値である。この値はネットワーク管理者によって設定された後、OSPF プロトコルによって各ルーター間で交換され、AS(Autonomous System : 自律システム) 内の全てのルーターへ通知される。各ルーターは、通知されたこの重みの値を仮想的なリンク距離として見なし、経路計算を行う。図 3.1 に 7 つのノードと 11 のリンクから構成されるネットワークの例を示す。このネットワークにおいて、送信ノード  $s$  から宛先ノード  $t$  までの最短経路は赤色の線で示されている。この経路は 2 章で紹介したダイクストラ法によって求められ、予め経路テーブルへと記述される。到着するパケットは、この経路テ

ブルを参照して経路を選択し，この最短経路を通過して宛先へと伝送される。

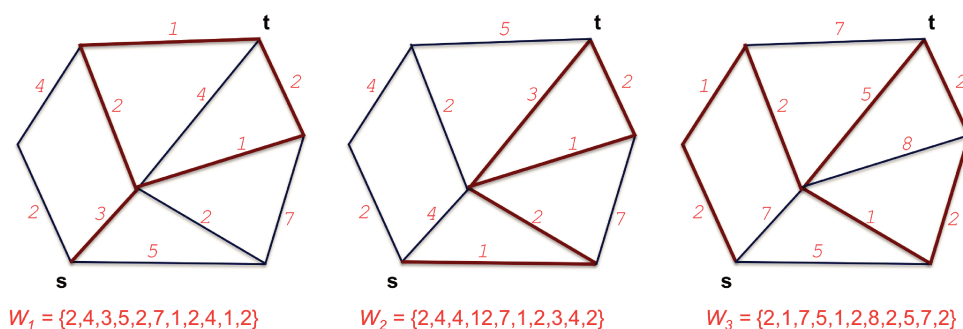


図 3.2: メトリック重み付けと最短経路の変化

次に，同じネットワーク上で，メトリックの重み付けを変化させた3つのパターンを図 3.2 に示した．各リンクに割り当てたメトリックの重み付けを配列  $W$  で表し，3つのパターンの  $W$  ( $W_1$ ,  $W_2$  および  $W_3$ ) に対する  $s$  から  $t$  までの最短経路が，それぞれの赤色の線である．ここで，OSPF では同一の距離を持つ経路があれば，ECMP(Equal Cost Multi Path：同一コスト複数パス)として扱い，トラヒックを等分に振り分ける．メトリックの適切な重み付けにより，ECMPを促進することができれば，負荷分散による輻輳の抑制が達成される。

実際には，全てのノードが送信ノードとなり，宛先ノードとなり得る事に注意してほしい．即ち，メトリック重み付けの変更は，全てのトラヒックの経路を変更する．メトリックが大きくなるリンクは，最短経路となりにくく，経路として選択されにくいという原理から，リンク帯域の逆数に基づいた各メトリックの重み付けも提案されているが[2]，しかし，もしノード間で交換されるトラヒックが予測できるならば，そのトラヒックに追従するような重み付けを計算する事が最適だろ

う。このような計算を OSPF 重み付け計算 (OSPF Weight Setting) と呼び、この問題は NP 困難に分類される。ネットワークの効率的なリソースの利用のために、この OSPF 重み計算を解き、ネットワーク全体の混雑を低減することは、ISP にとって重要な仕事である。

ここで、OSPF 重み付け問題を具体的に定義付ける。ネットワークは数学的なモデルである有向グラフ  $G = (N, E)$  として表す。ここで  $N$  はネットワーク上のノード数であり、 $E$  はエッジ数である。ノードはグラフ中の頂点であり、ルーターを表す。またエッジとはグラフ中にて方向を持った辺であり、実際のネットワークでは双方向リンクの片方を表す (すなわち、リンク数は  $E/2$  である)。一方で、ノード間で交換されるトラヒックは2章にてモデリングされたように、 $N \times N$  の行列として表す事ができる。グラフ  $G = (N, E)$  に表されるネットワークに対し、各ノード間に流れるトラヒックの要求が行列  $D$  によって与えられている場合、メトリックの重み配列  $W$  に対する各エッジ  $e$  の混雑率  $f_e/C_e$  を一意に決定することができる。(ここで、 $f_e$  はエッジ  $e$  に流れるトラヒックフローの合計であり、 $C_e$  はエッジ  $e$  のリンク帯域容量である) そして [3] によれば、最も混んだリンクの混雑率  $L = \max_{e \in E}(f_e/C_e)$  を最小化することが、ネットワークの利用効率を最大化する。以上より、与えられたトラヒック行列  $D$  に対して、 $L$  を最小化してネットワーク利用効率を最大化する重みセットを求める問題として、OSPF 重み付け問題を定義できる。

実用的な OSPF ネットワークでは、重み最適化計算に与えられた計算時間は有限である。ネットワークのトラヒックの要求が変化すると、その時点での重みセットは最適では無くなり、新たなトラヒック要求のための最適値の計算が必要となる。すなわち、時々刻々変化するトラヒック要求に対応するために、動的な重みセットの最適化計算が必要となる。また、トラヒック要求に応じて光パスのセットアップを変更するような光パスネットワークにおいても、同様の最適化が用いられる。そ

の場合、仮想ネットワークに対する OSPF の重みが計算され、その値に基づいて仮想ネットワークの変更が行われなければならない。このようなマルチレイヤーでの反復的な計算においても、短時間での計算が求められる。

OSPF-WS を解くためのアルゴリズムは、いくらかの研究者により研究された [4]。これらは大きく分けて、ILP(Integer Linear Programming : 線形計画法) により解く方法と、ヒューリスティックサーチアルゴリズムによって解く方法である。前者の ILP では、まずネットワーク最適化問題である OSPF-WS を、数学的 ILP モデルとして線形式に表現し、次に ILP ソルバーがこれを解く。ILP を用いた手法は、最適解を求めるようにデザインされたアルゴリズムであり、十分な時間をかけることにより非常に優れた解を出す。この手法は、十分な計算時間がある静的なアプリケーションでは一般的な方法である。しかし、ILP を用いて OSPF-WS を解いた場合、初期解を得られるまでの時間が非常に長いことが欠点として挙げられる。そのため、十分な計算時間がない場合では、解を一つも導出することができない可能性がある。よって、動的な OSPF-WS においては実用的ではない。

一方のヒューリスティックサーチは、最適解の導出を諦める代わりに、経験的な探索ベースの解決テクニックにより、短時間で十分に利用可能な解を導出する手法である。この方式では解の精度は保証されないが、計算中のいつでも、その時点までに求めた最適な解をいつでも取り出すことができる。時間的な制約に対して柔軟な手法であり、実用的な動的な OSPF-WS の解決には好適である。Ericsson ら [5] は、メタヒューリスティックの一つである GA (Genetic Algorithm : 遺伝的アルゴリズム) を利用し、OSPF-WS の解を求める手法を提案した。この手法によって求めた OSPF-WS の解は、大きめのネットワークに対して十分な時間を与えた場合、ILP と同程度か少し劣る解を導出する。しかしながら、大きなネットワークである場合では解が収束するまでに時間がかかってしまい、良い解を得ることができない。



アルゴリズムが十分に効率的である場合、計算時間を短くするための残された方法は、計算処理を行うプロセッサの高性能化に期待する事である。しかしながら、近年ではCPUのクロック性能は廃熱処理とともに頭打ちとなり、3GHzを前後に限界を迎えている。そのためCPUベンダーの多くは、クロック性能をそのままに計算コア数の増大へと方針を移した[6]。現在では、コア数が10個以上のマルチコアCPUも珍しくなく、今後もCPUコア数は増加する傾向にあると考えられる。このマルチコアCPUの性能をフルに引き出し、高い計算パフォーマンスを得るためには、並列処理アルゴリズムが必要となる。現在、OSPF-WSを解くアルゴリズムにおいて、並列処理により行った研究はまだ報告されていない。そこで、本研究ではOSPF-WSをマルチコア上で並列処理計算により高速に解くことに注目した。特に、先に述べた遺伝的アルゴリズムは、内部で行う反復的な処理が本質的に独立しており、並列処理に対し好適であった。しかし、これらの独立した処理だけを単純に並列化しただけではコア数に乗じたパフォーマンスを得る事ができない。そこで、単に並列化するだけでなく、独立に動作する各コアが矛盾しないように、同期により正しい計算順序を守り、且つ少ない待ち時間で効率的に処理させるアルゴリズムを提案した。

本稿では、OSPFネットワークにおける動的な重み付けによるトラフィックエンジニアリングのために、マルチコアCPU上での並列化によるOSPF-WSのより高速な解の導出方法について注目した。[5]の提案するGAによるOSPF-WS解決のアルゴリズムを基本とし、マルチコアCPU環境において高速に解くアルゴリズムを提案した。これらの並列化されたアルゴリズムは、16コアのIntel Xeon CPUにより実行され、シングルコアに比べておよそ13倍近く高速化される事を示した。そしてこの高速化により、大きなネットワークにおける重み付けにおいても実用的に解を求めることが可能となった。

### 3.3 OSPF 重み付け最適化の各方式

#### 3.3.1 ILP による OSPF 重み付け最適化

OSPF-WS を解く方法の一つは、ILP を用いる方法である。ILP は最適化計算のための数学的手法の一つで、与えられた制限下において、目的関数と呼ばれる 1 次式を最大化（または最小化）する値を求める方法である。OSPF-WS を ILP で解く為にはまず、OSPF-WS を一つの目的関数といくつかの制約条件により表さなければならない。この制約条件は整数を含む線形関数の等式もしくは不等式だけで構成される必要がある。

ILP によって OSPF-WS を解く手法として、Pioro ら [7] の手法が知られている。この研究では大きな数  $M$  を用いて、OSPF-WS の問題を次のような ILP 制約式として表した。OSPF ネットワークは有効グラフ  $G = (V, E)$  として表され、ノードがルーターを、エッジはリンクを表す。エッジ  $e \in E$  は  $c_e$  の帯域キャパシティを持つ。トラフィック要求  $D$  は行列として与えられ、その要素  $D_{st}$  は始点ノード  $s \in V$  から終点ノード  $t \in V$  に対する値を表す。また、 $V_d \subseteq V$  は終点ノードの集合である。

決定変数:

$f_e^t$  : エッジ  $e$  を通過する終点ノード  $t$  宛のトラフィックフローサイズ

$x_e^t$  : フロー  $f_e^t$  が 0 以上であるかどうかのバイナリ

$f_v^t$  : ノード  $v$  上におけるフロー分割のための仮想的なフローサイズ

$w_e$  : エッジ  $e$  における重みの値

$d_v^t$  : ノード  $v$  から終点ノード  $t$  までの最短距離

$L$  : 全エッジにおけるリンク負荷の最大値

フロー保存:

$$\sum_{e:e=(-,t)} f_e^t = \sum_{v \in V} D_{vt} \quad \forall t \in V_d \quad (3.1)$$

$$\sum_{e:e=(-,v)} f_e^t - \sum_{e:e=(v,-)} f_e^t = -D_{vt} \quad \forall v \in V \setminus \{t\} \quad \forall t \in V_d \quad (3.2)$$

制約式 (3.1), (3.2) は, 各ノードにおけるフローの保存を実現する. それぞれ, 終点ノードにおいてフローが終端すること, また通過ノードでフローが失われないということが表される.

フロー分割:

$$Mx_e^t - f_e^t \geq 0 \quad \forall e \in E, \quad t \in V_d \quad (3.3)$$

$$Mf_e^t - x_e^t \geq 0 \quad \forall e \in E, \quad t \in V_d \quad (3.4)$$

$$f_v^t - f_e^t \geq 0 \quad \forall e = (v, -) \in E, \quad \forall v \in V \setminus \{t\}, \quad \forall t \in V_d \quad (3.5)$$

$$f_v^t - f_e^t \leq M(1 - x_e^t) \quad \forall e = (v, -) \in E, \quad \forall v \in V \setminus \{t\}, \quad \forall t \in V_d \quad (3.6)$$

大きな数  $M$  を導入することによって, バイナリ  $x_e^t$  と  $f_e^t$  の関係が対応付けられる. (この  $M$  は  $\sum_{v \in V} D_{vt}$  として置き換えが可能である) 制約式 (3.3) および (3.4) により,  $x_e^t$  と  $f_e^t$  は同時にゼロとなるか, 同時に非ゼロとなるかのいずれかとなる. よって定義の通り,  $x_e^t$  はフロー  $f_e^t$  の有無を表す. 残る式中で用いられる  $f_v^t$  は, ノード  $v$  におけるフロー分割のための仮の値である. 制約式 (3.5) および (3.6) によって, ( $e = (v, -)$ ) となるエッジ  $e$  (すなわち, ノード  $v$  から隣接ノードへのリンク) 上のフロー  $f_e^t$  は, 0 もしくは  $f_v^t$  のいずれかの値しか取る事ができず, 結果 ECMP が実現される.

最短距離ルーチング :

$$d_v^t + w_e - d_u^t \geq 0 \quad \forall e = (u, v) \in E, \quad \forall t \in V_d \quad (3.7)$$

$$d_v^t + w_e - d_u^t \leq M(1 - x_e^t) \quad \forall e = (u, v) \in E, \quad \forall t \in V_d \quad (3.8)$$

$$M(d_v^t + w_e - d_u^t) \geq 1 - x_e^t \quad \forall e = (u, v) \in E, \quad \forall t \in V_d \quad (3.9)$$

制約式 (3.7), (3.8), (3.9) の作用は少し複雑である. もしエッジ  $e$  上に終点ノード  $t$  へのフローが流れていれば, すなわち  $x_e^t = 1$  であれば, 制約式 (3.7), (3.8) によって  $d_v^t + w_e = d_u^t$  となる. 各隣接ノードがこのエッジ  $e$  に対する関係性を持つ事によって,  $d_u^t$  はパス上の  $w_e$  の合計値となる. 一方で,  $x_e^t = 0$  となるパス上にないエッジ  $e$  に関して, 制約式 (3.9) により  $d_v^t + w_e$  は  $d_u^t$  よりも大きな値となるため, 最短経路ではないということになる. 以上のようにして, もしフローが存在し,  $x_e^t = 1$  となる場合にエッジ  $e$  が最短経路上に存在する.

リンク混雑率 :

$$\sum_{t \in V_d} f_e^t \leq Lc_e \quad \forall e \in E, \quad \forall t \in V_d \quad (3.10)$$

最後の制約式 (3.10) によって, 全てのエッジにおけるリンク混雑率の最大値は  $L$  以下でなければならない. 上記の式を用い, ILP は以下のように表現される.

*minimize*  $L$

$$s.t. \quad x_e^t = \{0, 1\}, \quad f_e^t \geq 0, \quad f_v^t \geq 0, \quad w_e \in \mathbb{N}, \quad d_v^t \in \mathbb{N}, \quad L \geq 0$$

上記の式によって表された ILP 問題は, ILP ソルバーと呼ばれるソフトウェアによりコンピュータ上で解かれる. CPLEX は IBM によって開発された優れた ILP ソルバーであり, 市販のソルバーとして最も知られているもの一つとして, [7] においても用いられている. CPLEX では上記の制約条件式と目的関数を入力ファイル

として読み込み、計算を開始する。これらの複数の制約条件式は解析され、解の探索が行われる。多くの場合、CPLEX は後に説明されるヒューリスティックサーチアルゴリズムによりも優れた解を導出する。

ILP ソルバーが求める解は、重みセットだけではなく、式中に登場する全ての変数の組み合わせである。すなわち、探索対象の解空間はいずれかの制約条件を満たさない実行不可能な解を含む。そのため、ILP のソルバーは目的関数の最小化を試みつつ、最初の実行可能な初期解を見つけるまで、意味を持つ重みセットを持たない事になる。この初期解を得るまでの時間は、計算量と共に爆発的に長くなる。特に、大きなサイズのネットワークでは、実用的な計算時間内に一つも解が見つからない可能性がある。その場合、動的重み付けのようなアプリケーションには利用することができない。

### 3.3.2 GA による OSPF 重み付け最適化

一方のヒューリスティックサーチによる手法は、最適解の導出をあきらめる代わりに、経験則からより良い解を求めるというアプローチである。この手法では解の精度は保証されないが、ある程度最適解に近い解を短時間に得ることが出来る。ヒューリスティックサーチによる OSPF-WS の解決に関する研究として、メタヒューリスティクスを用いた手法が報告されている [4]。メタヒューリスティクスとは具体的なアルゴリズムを含まない概念的な解法の事である。OSPF-WS へのメタヒューリスティクスの適用ではローカルサーチ [8]、焼き鈍し法 [9] が挙げられる。そして、Ericsson ら [5] は GA を用いて OSPF-WS を解いた。この GA は、アルゴリズム内部で独立した複数の処理を行う必要性があり、とくに並列コンピューティングに好適である。

GA[10] は Holland によって提案されたメタヒューリスティックであり、ダーウィ

ンの進化論に基づいて組み合わせ最適化問題を解くことができる。GA では解空間の各候補解（個と呼ばれる）を染色体として表現し、様々な染色体を用意して集団を形成する。集団内の染色体を評価し、その評価値に基づいて継承、突然変異、選択、交叉などの進化過程の模倣処理を反復的に行う。各反復を世代と呼び、世代毎に集団は進化する。染色体を評価する値を適応度と呼び、進化過程では、適応度が優れてそうな染色体を選んでは、次の世代にその遺伝子が残るように行われる。結果、世代毎に集団の適応度は改善される。ここで、計算機上でデータ型として表される染色体をどのようにエンコーディングするか、また何を適応度としてどのように求めるか、そしてどのように進化過程を行うかの具体的な動作はアルゴリズムに依存する。これらはアルゴリズムの性質を大きく左右する重要な要素である。

Ericsson ら [5] の GA は、OSPF-WS における重みセットそのものを遺伝子としてエンコーディングし、また最大リンク混雑度である  $L$  を適応度とした。初期遺伝子は全て乱数によって与えられ、世代あたりの集団サイズを  $P$  で一定とする。各世代では、全ての遺伝子が適応度を評価された後、進化過程によって次の世代が生成される。ここで、それぞれを評価ルーチン、および進化ルーチンと呼ぶ。評価ルーチンでは、集団内の染色体のうち、適応度が優れた一部をエリートと呼び、これらはそのままグループ A として残す。残りの染色体については非エリートとし、その一部は交叉によって生成した染色体と入れ替え（グループ B）、残りはランダム発生した染色体と入れ替えられる（グループ C）各グループのサイズを  $A, B, C$  と表した場合、 $P = A + B + C$  である。（同様にエリートの数は  $A$  である）

この遺伝的アルゴリズムの最も重要な処理は、図 3.3 に表される交叉による染色体グループ B の生成である。一度の交叉では、エリート染色体から選ばれた一つと、非エリート染色体から選ばれた一つから、新しい一つの染色体が生成される。生成される染色体の重みセットの各エッジの重みの値は、基となるエリートおよび非エ

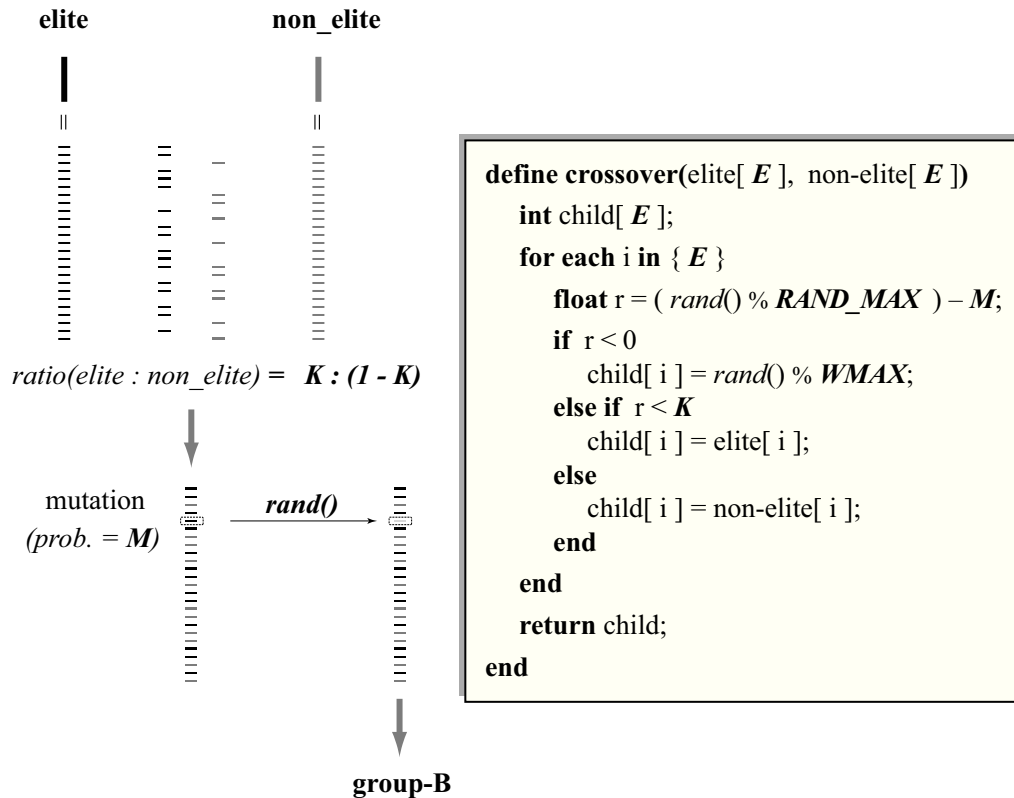


図 3.3: 交叉

リート染色体のそれぞれからどちらかの値をとる。その確率はパラメータ  $K$  によりコントロールされ、 $K$  の確率でエリート染色体上の重み配列  $W$  が、 $(1 - K)$  の確率で非エリート染色体上の値が選ばれる。さらに、パラメータ  $M$  を確率として、非常に稀に突然変異として乱数値と交換される。重みセット上の各重みに対する  $M$  と  $K$  を用いた操作において、乱数の発生を最小に抑えるコードは図 3.3 の右側に表される。この操作によって生成された染色体は、各段階で優秀な解であるグループ A に類似した解候補であり、既存のエリート染色体よりも優れた遺伝子を持つ可能性を含む。そして、次の世代で適応度が計算され、もしも優れていれば次のエリート

染色体としてグループ A に残されてゆく。このようにして、世代ごとに集団全体の遺伝子は進化していく。

## 3.4 マルチコア CPU による

### OSPF 重み付け最適化の並列計算

ここまで、OSPF 重み付け最適化の手法として、ILP によるものと GA を行うものをそれぞれ紹介した。ILP を用いたものでは、高い最適度の解（すなわち、より小さな目的関数  $L$ ）となるメトリック重み配列が得られるが、計算時間が非常に長い。一方で、GA はヒューリスティックな解法により、比較的短時間で最適解に近い解を導出する。特に、GA は本質的に、それぞれが独立した反復的な処理を含む為に、並列処理には好適である。

3.4 に GA における概念を示す。図中の  $W$  の評価を行う処理は、本質的に独立であり、容易に並列化が可能である。しかしながらこの方式では、コアの待ち時間が多く、特にコア数の多い環境において、効率的な並列処理が行われない。そこで、提案方式では、データの依存関係を解決した効率的な並列アルゴリズムを実現した。本章では、この従来方式および提案方式それぞれについて説明する。両方式において、並列化は OpenMP を利用して実現される。また、途中で必要となる乱数発生のための `random()` 関数の使用には注意しなければならない。これらの関数は乱数テーブルを参照して新たな乱数を発生させるため、このテーブルを共用してはならない。乱数テーブルをコア数分だけ用意し、各スレッドがそれぞれのテーブルを参照するように拡張しておく必要がある。



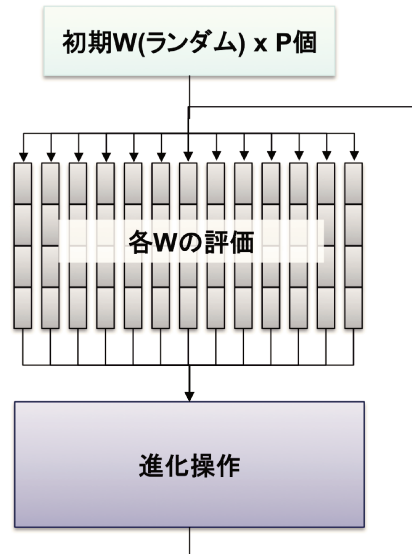


図 3.4: GA の基本的なアルゴリズム

### 3.4.1 従来方式：単純な並列化による OSPF 重み付けの最適化計算

まずはじめに [5] のアルゴリズムを単純に並列化した従来方式について、図 3.5 を用いて説明する。初期世代としてランダムに与えられた全ての染色体は、`evaluation()` 関数により評価される。`evaluation()` 関数の内部では、仮想的なネットワークにトラフィックフローを割り当てて適応度を計算する。重み配列  $W$  を実際のネットワークに割り当て、ダイクストラのアルゴリズム [14] を実行し、トラフィック要求のあるノード間の最短経路を計算する。そしてその最短経路上を通過するトラフィックからリンクの使用帯域を割り当て、その合計をリンク容量で割る事によって全リンクの混雑度  $L$  を求め適応度として返す。この計算を、グループ B およびグループ C の全ての染色体に対して実行し、全ての  $L$  を求める一連の処理が評価ルーチンである。こ

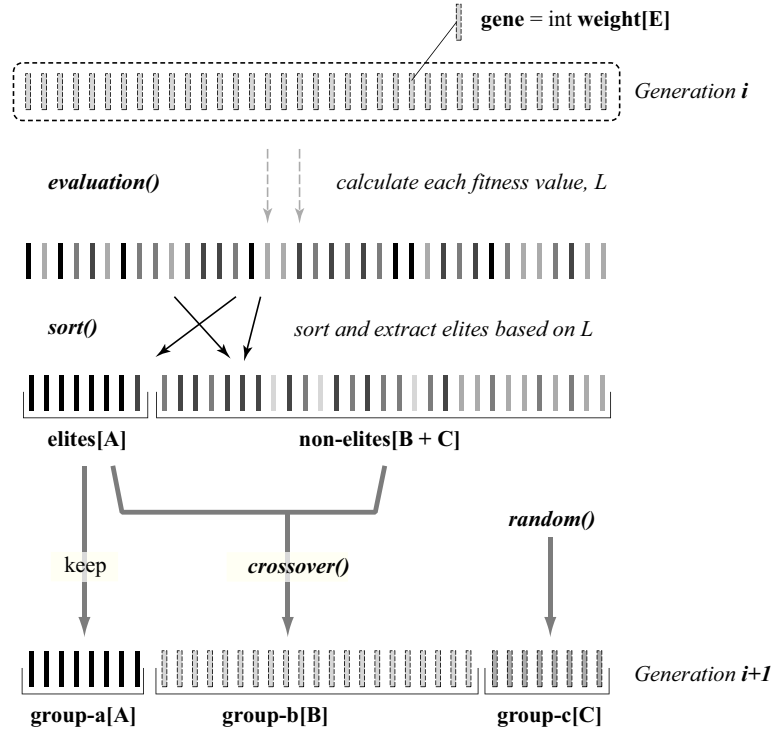


図 3.5: Ericsson の GA による OSPF 重み付け

ここで、染色体すなわち重みセットを評価する `evaluation()` 関数では、入力である各染色体データと、出力である適応度  $L$  のデータがそれぞれ独立である。そこで、各染色体に対する関数の処理を1つのタスクとし、これを各コアに分散して割り当て、並列に処理を行う。割当は動的に行われる、即ちタスクを終えたスレッドから順に次のタスクを処理する。全てのタスクが完了した時点で、各コアはスレッド同期を行い他の全てのスレッドの処理が完了まで待機する。これにより、残る進化ルーチンの処理を除き、評価ルーチンの並列化が完成する。

一方の次の進化ルーチンはデータに依存関係を持つため、1つのコア上で実行されなければならない。全集団を適応度に基づいてクイックソートし、エリート染

染色体と非エリート染色体を分類する。ここで、ソートの際データの書き換え時間を最小限に抑えるため、各データは、染色体データのインスタンスに対するポインタと、適応度  $L$  を組み合わせた構造体の配列として管理されてソートされる。この配列の先頭  $A$  個の要素はグループ  $A$  の染色体を指し、次の  $B$  個はグループ  $B$  を、残りの  $C$  はグループ  $C$  を指す。ソートされた後に、次の世代のグループ  $B$  の各遺伝子が交叉によって生成される。1回の交叉では、エリート染色体と非エリート染色体が乱数によって選択され、`crossover()` 関数によって一つの新たな遺伝子が生成される。ここで、全ての遺伝子が生成において、現世代の染色体は参照される可能性があるため、そのデータは書き換える事ができない。そのため、新しい遺伝子はバッファ上に作成され、全ての `crossover()` 関数の実行が終わった後に、ポインタがスワップされる。次の世代のためのグループ  $B$  の染色体が全ての交叉によって生成された後、今度はグループ  $C$  の染色体が生成される。グループ  $C$  の染色体は初期解と同様、ランダムな値によって置き換えられる。置き換えが完了したところで、その世代の進化ルーチンは完了し、次の世代へと進む。

## 3.4.2 提案方式：データの依存関係を解決した効率的な並列処理

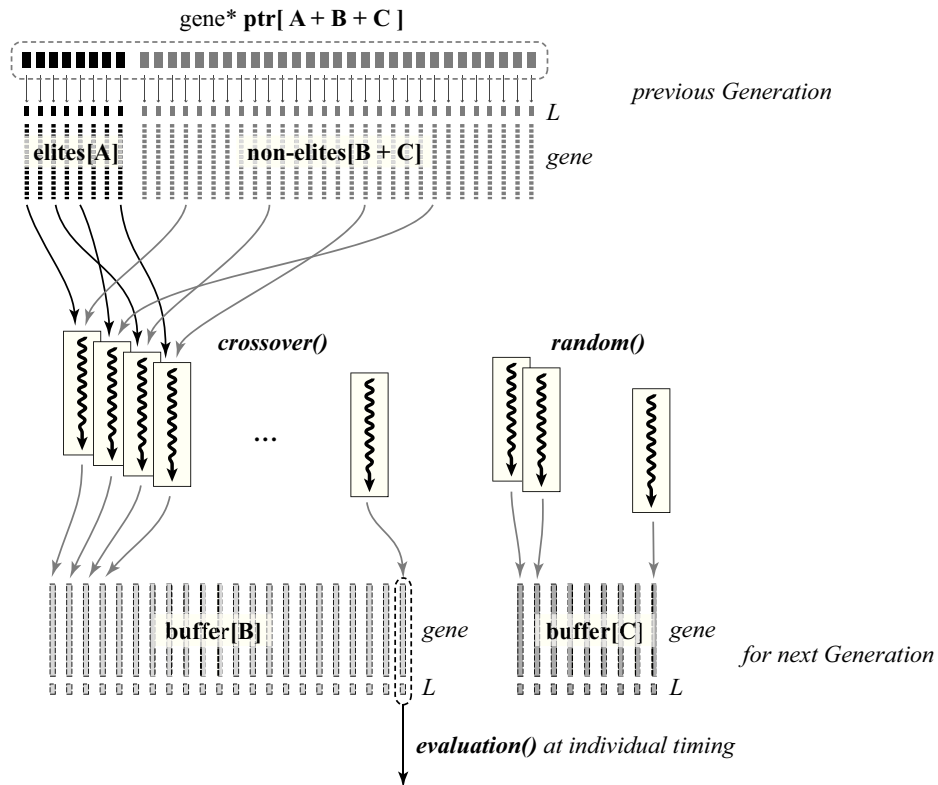


図 3.6: 提案方式のデータ構造

提案方式では、進化ルーチンを巻き込んで並列に実行するために、データの依存関係を解決した。進化ルーチンにおいて最初に行われるソートだが、クイックソートを使用した場合の処理時間は非常に短く、並列化を行わなくても十分なパフォーマンスが見込まれる。それどころか、データのコヒーレンスを考えた場合、同期を含む並列ソートではない方が良い。残る処理である、次の世代のグループ B のための **crossover()** 関数とグループ C に対する乱数による遺伝子の置き換えについての依

存性を考えよう。ここで提案方式では、図 3.6 のようにグループ C に対してもバッファを用意することとする。これによって、グループ C とグループ B のそれぞれの遺伝子の依存性は解決される。さらに、染色体のデータを管理する構造体において、重みセットのデータだけでなく、適応度  $L$  に対してもバッファを用意する。両データは構造体として管理するとともに、この構造体へのポインタを配列として管理する。これによって、`evaluation()` 関数における書き込み先に関しても、データ書き込み依存性を持たない。

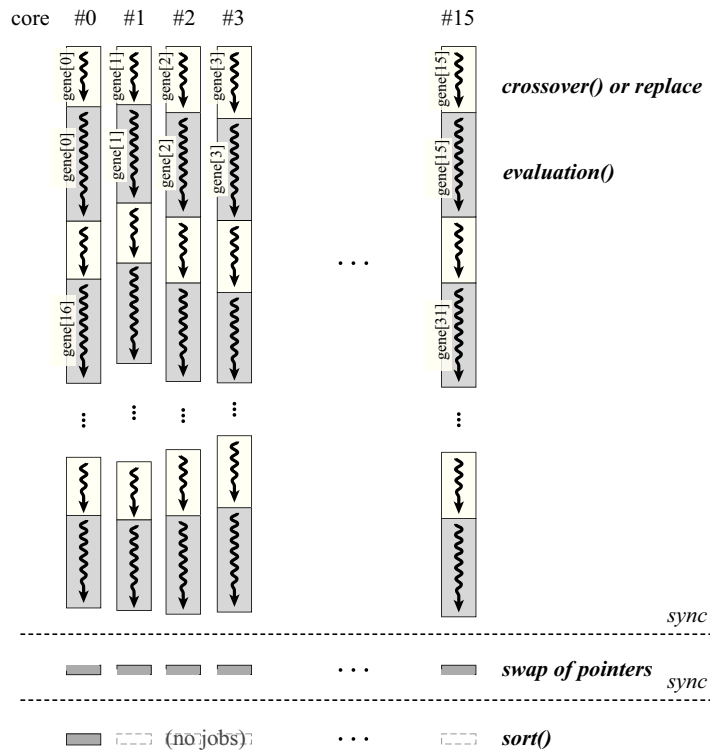


図 3.7: 提案方式による並列化

この依存性の解決により、次の世代を作成する `crossover()` 関数や、乱数による

遺伝子の置き換えの処理は、並列に実行している評価ルーチンの中で、`evaluation()` 関数の直前に実行する事ができる。最終的な処理の実行は図 3.7 のようになる。このアルゴリズムでは、コアの待ち時間が削減され、コア数に応じた効率的な処理が期待される。

## 3.5 性能評価

ここまでで説明した提案アルゴリズムによって、OSPF 重み付け最適化を行う GA が並列化された。本章では、この並列アルゴリズムのパフォーマンスを評価する。

まず、コンピュータ上で Waxman モデル [15] に基づきランダムなネットワークを発生させ、ランダムに選んだノードペア間にトラフィック要求を発生させる。このトラフィック要求を満たし  $L$  を最小化する OSPF 重み付け問題を定義し、重みセットを求めた。ILP による手法では [7] の式用い、CPLEX をソルバーとして利用する。一方の GA を用いた手法では各パラメータは  $A=300$ ,  $B=3000$ ,  $C=300$ ,  $K=0.5$ ,  $M=0.01$  とした。いずれの計算も Intel(R) Xeon(R) の CPU E5-2687W 3.10GHz 16core によって計算した。

### 3.5.1 並列化による計算時間の短縮

まず最初に、並列化によりどれくらい計算の速度が高速化されたかを評価した。本研究における並列化では [5] に対して解の求め方は変わっていない。故に、1つの世代を処理する時間をそのまま解を求める速度として評価する事が可能である。よって、世代あたりの処理時間を指標として評価を行った。10ノードのネットワークを10種類用意し、10ペアの始点ノードと終点ノードの間にトラフィックデマンドを発生させた。これらを100世代の反復だけ実行した処理時間を、5回の試行の平均値と

して測定し評価を行った。

表 3.1: 実行速度の差

	非並列	従来方式	提案方式
実行時間 (秒)	3.068	0.375	0.233
速度比	1.0	8.17	13.16
IPC	1.14	0.82	1.07
命令数 (x10 億命令)	13.25	16.63	13.58

表 3.1 に並列化によるスピードアップの結果を表す。速度比の値は、非並列での実行時間を 1.0 とした場合の比率の逆数であり、相対的な実行速度である。評価ルーチンのみを並列化した従来方式では、16 コアで 8.17 倍しか高速化されないのに対し、全て並列化した場合は 13.16 倍まで高速化される。この値は、アムダールの法則において計算すると 99 パーセント以上の並列化が達成された事を示す。コア数である 16 倍まで高速化されない理由としては、まずマルチコアにおいて同一データを共有する場合、コヒーレンシを保つためにキャッシュの有効性が低下する。これにより IPC (Instruction per cycle : サイクルあたりの実行命令数) が 1.14 から 1.07 へと低下しているのが確認できる。加えて、複数の並列セッションで処理を行う為に、同期やタスク割当などの命令を実行する必要がある。

### 3.5.2 重み付け問題での最適化計算結果の比較

今度は、実際に最適化している  $L$  の値として評価を行う。50 ノードのネットワークを 10 種類用意し、200 ペアのノード間へのトラフィック要求を発生させた。この問題を非並列での処理と、提案方式により並列処理で 60 秒間計算した場合の OSPF

重み付け最適化計算の結果を比較した。実験に用いたマシンは先ほどと同様であり、16 コア使用して並列に処理する。

表 3.2: 各ネットワークで 60 秒間に計算した混雑度  $L$  の値

Network	非並列	提案方式	非並列/提案
#0	0.2512	0.1538	1.63
#1	0.2268	0.1724	1.32
#2	0.2390	0.1982	1.21
#3	0.2710	0.1697	1.60
#4	0.2323	0.1403	1.66
#5	0.2467	0.1967	1.25
#6	0.2163	0.1262	1.71
#7	0.1893	0.1179	1.61
#8	0.2844	0.2333	1.22
#9	0.2779	0.2000	1.39
Average			1.46

表 3.2 に各方式で求めた  $L$  の値を示す。 $L$  の下限、即ち全空間での最適値はネットワークによって異なるため、それぞれのネットワーク毎に値を表した。提案方式では、非並列と比べ 1.2 から 1.7 倍まで  $L$  が小さい。また、図 3.8 に、時間に対する  $L$  の値の変化の例を示した。横軸の時間は対数であることに注意してほしい。



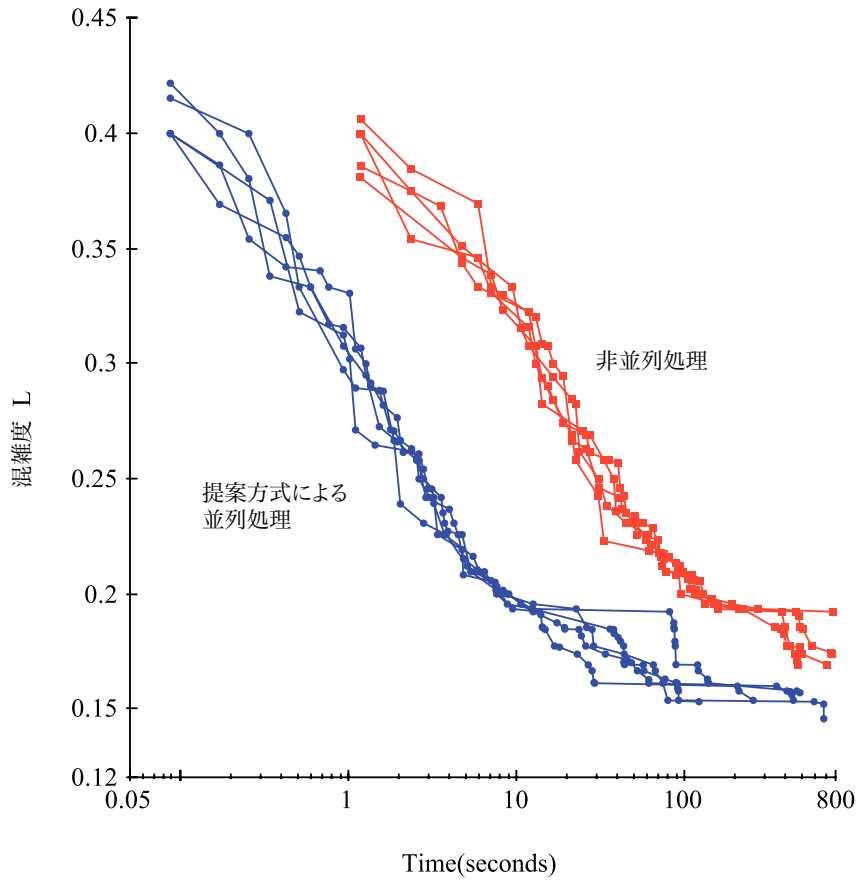


図 3.8: 非並列処理と提案方式による並列処理での混雑度  $L$  の比較

### 3.5.3 他のアルゴリズムとの最適化計算結果の比較

今度は，提案方式の OSPF 重み付け最適化が，他の ILP や，固定値による設定と比べて効果的かを比較した．ネットワークサイズは 20 ノードであり，800 秒間である．表 3.3 は提案方式 (並列処理 GA)，ILP(CPLEX) を用いて計算した結果に加え，全ての重みを 1 として最小ホップルーチングを行った場合 (minhop)，また Cisco が推奨する帯域の逆数とした場合 (cisco) の結果をそれぞれ表す．

表 3.3: 他のスキームとの混雑度 L の値の比較

network	Minhop	Cisco	ILP	Prop.	ILP/Prop.
#0	0.0833	0.3333	–	0.0244	–
#1	0.0690	0.1563	0.0500	0.0500	1.00
#2	0.0682	0.1667	–	0.0235	–
#3	0.0857	0.1333	0.0333	0.0333	1.00
#4	0.0761	0.0870	0.0500	0.0500	1.00
#5	0.1053	0.1333	0.0187	0.0208	0.90
#6	0.0714	0.2353	0.0200	0.0250	0.80
#7	0.1023	0.0909	–	0.0215	–
#8	0.0606	0.1429	0.0257	0.0262	0.98
#9	0.1071	0.1875	0.0714	0.0313	2.28

また一例としてネットワーク #2 の結果を図 3.9 にも表した．ノード数が小さいため，GA による方式は比較的早い時間で収束する．図 3.9 の例では 0.1 秒程度であり，同サイズの他のネットワークでもそれほど大きな差はない．一方の ILP は，GA よりも優れた解を求める事が多いが，ネットワーク #0, #2, #7 のように一つも解を

求められなかったケースもある。これらは表中でハイフンで表された、初期解の時間が 800sec を超えたことを意味する。これがさらに大きなネットワークであれば、解を出す事自体が困難であることは明白である。その他の固定値で重み付けを行った最小ホップルーティングや帯域の逆数に比べれば十分であり、提案の GA は ILP による手法と比べ同程度の最適化が可能であると言える。

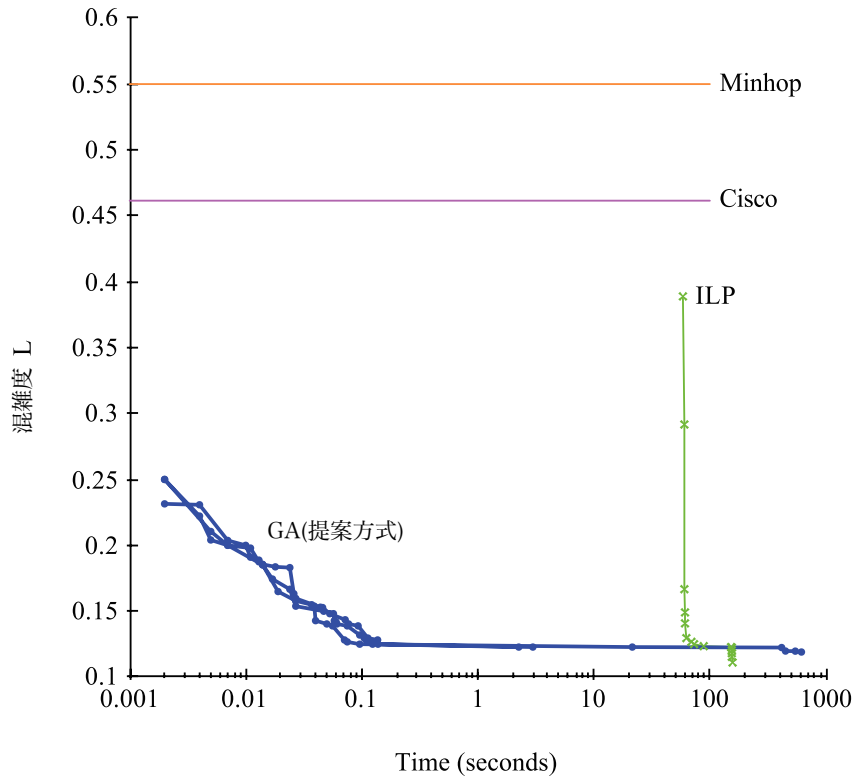


図 3.9: 提案 GA および他のスキームを用いた混雑度 L の最適化の結果

## 3.6 結論

IP ベースの TE では、OSPF のメトリック重み付け最適化によって TE を行う。しかし、この最適化問題は NP 困難であり、その計算量は非常に大きい。ILP による計算方式は十分に優れた解を見つける事ができるが、初期解までの時間が非常に長いため、計算時間が限られた状況では利用できない。一方のヒューリスティックな方式として、GA による計算は非常に有効である。それでもなお大きなサイズのネットワークに対しては、収束までに非常に長い計算時間が必要となる。この計算時間を短くする為の方法として、近年ではマルチコアによる並列処理が注目されている。GA は本質的にデータ依存が少なく、独立した反復処理を含む為に並列処理に好適であった。しかしながら、単に反復処理を並列化するだけでは、コアの待ち時間が発生し、特にコア数が多い場合に置いて非効率な計算となる。そこで、本研究では、データの依存性を解決し、効率的な並列アルゴリズムを提案した。両方式は 16 コア CPU によって計算され、従来の方式が 8 倍程度しか高速化しないのに対し、13 倍近い高速化を得た。これにより、今後の大規模並列化時代に向けた、最適化計算の規模拡張性の可能性を示した。

## 参考文献

- [1] J. Moy, "OSPF Version 2," IETF Proposed Standard, RFC 2328, April 1998.
- [2] T.M. Thomas II. "OSPF Network Design Solutions." Cisco Press, 1998.
- [3] Eiji Oki and Ayako Iwaki, "Load-Balanced IP Routing Scheme Based on Shortest Paths in Hose Model," IEEE Transactions on Communications, Volume 58, Issue 7, Page 2088-2096, July 2010
- [4] Ghazala, A.A., et al., "A Survey for Open Shortest Path First Weight Setting (OSPFWS) Problem," Proceedings of International Conference on Information Security and Assurance(ISA), April 2008.
- [5] M. Ericsson, et al., "A Genetic Algorithm For The Weight Setting Problem In Ospf Routing," Journal of Combinatorial Optimization, Volume 6, Issue 3, Volume 6, Page 299-333, September 2002.
- [6] D.A. Patterson, J.L. Hennessy, "Computer Organization and Design," Fourth Edition, Morgan Kaufmann, November 2011.
- [7] M. Pioro, et al., "On open shortest path first related network optimization problems," Journal of Combinatorial Optimization, Volume 48, Page 201-223, Volume 48, April 2002.

- [8] Bernard Fortz and Mikkell Thorup, "Internet traffic engineering by optimizing OSPF weights," Proceedings of IEEE INFOCOM, Page 519-528, Volume 2, March 2000.
- [9] Mohammed H. Sqalli, et al., "An Enhanced Estimator to Multi-objective OSPF Weight Setting Problem", Network Operations and Management Symposium (NOMS), Page 240-247, April 2006.
- [10] J.H. Holland. "Adaptation in Natural and Artificial Systems," MIT Press, 1975.
- [11] Ting Lu, et al., "Genetic Algorithm for Energy-Efficient QoS Multicast Routing," IEEE Communications Letters, Volume 17, Issue 1, Page 31-34, January 2013.
- [12] Zuqing Zhu, et al., "Using Genetic Algorithm to Optimize Mixed Placement of 1R/2R/3R Regenerators in Translucent Lightpaths for Energy-Efficient Design," IEEE Communications Letters, Volume 16, Issue 2, Page 262-264, February 2012.
- [13] Gene M. Amdah, "Validity of the Single Processor Approach to Achieving Large-Scale Computing Capabilities," Proceeding of AFIPS Conference, Page 483-485, April 1967.
- [14] Dijkstra, E. W., "A note on two problems in connection with graphs," Numerische Mathematik, Page 269-271, 1959.
- [15] B.M.Waxman, "Routing of multipoint connections," IEEE Journal on Selected Areas in Communications, Volume 6, issue 9, Page 1617-1622, December 1988.

---

## Chapter 4

---

# GMPLS ベースのマルチメトリック トラフィックエンジニアリングにおける 制御システムの規模拡張性

### 4.1 概要

第3章では、IP ベースの TE における最適化計算に関する問題を取り扱った。IP ベースの TE ではメトリックの重み付けによりトラフィックを制御する一方、GMPLS では LSP によるトラフィックの制御を行う。LSP は動的かつ柔軟なトラフィックを制御を可能とした。これによって輻輳の解決だけでなく、アプリケーションのための QoS 制御や、物理レイヤーのパラメータを考慮した TE などに期待が寄せられている。しかし、これらのパラメータ多様化によるネットワークのマルチメトリック化は、頻繁なメトリック更新を伴う。そしてこのメトリック更新のメッセージのために、ネットワークが逼迫されることが問題となる。この問題に対し、本研究ではフラッディングに基づくシグナリングによるアプローチを提案した。提案方式では、ルーチングプロトコルを停止する事により、メトリック情報の配布を行わない。その代わりに、シグナリングメッセージをフラッディングし、最短経路を導出しながら LSP を確立する。この方式では、メッセージの発生頻度はメトリック更新頻度に依存しな

いため、メトリック更新が頻繁なネットワークにおいて大幅にメッセージの発生を抑える事ができる。これらはコンピュータシミュレーションにより、ブロック率を保ちながらメッセージが低減する事が検証された。またセットアップ時間などについても同様にシミュレーションに基づき考察された。

## 4.2 マルチメトリック環境でのパス確立

GMPLS (Generalized Multi-Protocol Label Switching) [1] はコネクションオリエンテッド通信の確立を自動化するフレームワークである。MPLS (MultiProtocol Label Switching) [2] が従来のレイヤーモデルにおけるパケット交換ネットワークだけをサポートするのに対し、GMPLS は WDM (Wavelength Division Multiplexing : 波長分割多重) ネットワークや、同期型デジタル通信方式である SONET/SDH (Synchronous Digital Hierarchy / Synchronous Optical Networking) ネットワーク、ATM (Asynchronous Transfer Mode : 非同期転送モード) ネットワークなどの様々なネットワーク、さらにはこれらの相互接続ネットワークなどの様々な多様化したネットワークをサポートする。GMPLS は既存のあらゆるネットワークを統合する次世代ネットワーク技術として期待されている。

GMPLS が確立するコネクションは LSP (Label Switched Path : ラベル交換パス) と呼ばれる。この LSP をオンデマンドに確立することによって GMPLS はデータトラヒックの流れを制御する。GMPLS における LSP 確立は、3つのステップからなる。(1) ルーチングプロトコルである OSPF-TE (Open Shortest Path First - Traffic Engineering) [3] によってそれぞれのリンクやノードの状態などのリソース情報をメトリック情報としてネットワーク全体へと通知する。(2) ソースノードのパス計算エンジンが、最新のメトリック情報を基に CSPF (Constraint Shortest Path First) ア



ルゴリズムを実行し、LSP 要求を満足させる経路を導出する。(3) シグナリングプロトコルである RSVP-TE (Resource Reservation Protocol - Traffic Engineering)[4] によってシグナリングを行い、経路上の全ノードのルーティングもしくはスイッチングステートを変更することにより、要求された LSP を確立する。LSP の要求には確立される LSP に対するいくつかのパラメータが含まれる。このパラメータは、必要な通信帯域などのカスタマーの要求だけでなく、TE (Traffic Engineering) による最小のリソース使用などネットワークプロバイダの要求も含まれる。両者の様々な LSP 要求に応じる為に、ルーティングプロトコル、パス計算エンジン、シグナリングプロトコルのそれぞれはメトリック情報多様化に対して備えなければならない。

現在の GMPLS プロトコル標準は、いくつかのネットワークメトリックをメトリック情報として使用できるよう対応している。TE メトリックと呼ばれる値は、TE を実現する為にネットワーク管理者によって設定された値である。8つに分類された未使用帯域の値は、優先度の異なる帯域保証に使用される。さらに、GMPLS は今後の新しいサービスや技術などに使用されるための新たなメトリックのために、拡張の余地を残している。その中の一つとして、最も標準化に近いものが PLI (Physical Layer Impairment : 物理信号劣化) 評価である [5]。この拡張では、PMD(Polarization Mode Dispersion : 偏波モード分散) や CD (Chromatic Dispersion : 色分散), ASE(Amplifier Spontaneous Emission : 増幅自然放出光) ノイズなどの光信号の情報をメトリック情報として扱う事により、光信号の信号劣化を考慮した経路選択が可能となる。PLI のための拡張に関する標準化は現在、IETF により行われている。[6] [7] また [8] では、リンク損失率をメトリックとして考慮した信頼性の高い LSP 確立のための拡張に向けた研究がされている。GMPLS が対応すべきメトリック情報は、ネットワークに関するメトリックだけではない。近年のいくつかの研究では、電力消費量を考慮したエコロジーネットワークとして、電力量

をメトリックとする拡張が検討されている [9] [10]. ユビキタスネットワーク環境では、様々なサービスやアプリケーションに関する情報が LSP 確立において考慮される [11]. また、無線の分野ではモバイルバックホールネットワークにおける一つの基地局あたりの端末数を考慮した LSP 確立、CDN (Contents Delivery Network : コンテンツ配信ネットワーク) ではコンテンツサーバーあたりのユーザー数、その他の要素など様々なサービスに関係するメトリックが GMPLS によって対応される必要がある。

現在の GMPLS のパス確立手法では、経路確立時に始点ノードが CSPF 計算を行う。この CSPF 計算においては、ネットワーク全体のメトリック情報が使用される。必要なメトリック情報は確立中の LSP が要求とするものだけで十分だが、次に要求される LSP がどのメトリック情報を必要とするかは予測する事ができない。故に、始点ノードとなる可能性のある全てのノードは、ネットワーク中の全ノードについて、サポートされる全てのメトリック情報を知っていなければならないことになる。さらに、そのメトリックは常に最新状態でなければならない。つまり、一度リソースの状態が変更されると、ネットワーク上の全てのノードが持つデータベースを更新するために、ネットワーク全体に更新メッセージを送信する必要がある。結果的に、たった一度のリソースの状態の変更がメッセージのフラッディングを起こし、さらに全てのネットワークノードはこのメッセージを処理することに追われる。GMPLS が多くのネットワークサービスに対応しメトリック情報が増えるにつれ、データベース更新のためにフラッディングするメッセージの数は莫大となり、メッセージ処理が追いつかないかもしれない。このような事態を避ける為に、サポートするメトリック情報が増えても、更新メッセージがネットワークを埋め尽くさないような新たな手法が要求される。

本章では、メトリックの多様化に対応し、最小のオーバヘッドによる非常に効

率的な LSP 確立手法を提案する。提案手法は拡張シグナリングプロトコルである FB-RSVP-TE (Flooding Based RSVP-TE : フラッディングに基づいた RSVP-TE) プロトコルによって CSPF 計算を実現する。この確立手法では、ルーティングプロトコルを必要とせず、メトリック情報が変更されるたびに更新メッセージをネットワーク上へフラッディングする必要が無い。そのため、メトリック情報の更新が頻繁に行われようと、ネットワーク上を通過する制御メッセージは少ないままである。提案手法は、メトリックの変更に対する優れた柔軟性だけでなく、従来の手法と同程度の低いブロック率を発揮する。

FB-RSVP-TE はフラッディングに基づく経路探索であり、無線ネットワークの分野では DSR (Dynamic Source Routing : 動的ソースルーティング) [12] として研究されている。DSR の目的は、ネットワークトポロジが固定されていない無線ネットワークにおいて経路を探索し通信を確立することである。そのため、DSR は最小ホップで経路探索を行うが、様々なメトリックを最小とする経路探索は行わない。そこで、我々はこのアプローチに基づいた二つのメカニズムにより、FB-RSVP-TE を実現した。これらのメカニズムの詳細は、本章にて紹介される。また、FB-RSVP-TE の性能は、導出した経路およびメッセージ更新のコストの両面で評価した。

既存の GMPLS における LSP 確立手法に関する問題点は 4.3 節において説明される。また、提案手法は 4.4 節において紹介される。これらの両手法は、4.5 節において評価された。そして、4.6 節にて本章が結論づけられる。

### 4.3 従来手法

図 4.1 に GMPLS ネットワークの概要を示す。このネットワークは論理的に二つのプレーン (plane, 面) に分離されている。一つは制御プレーンであり、もう一つ

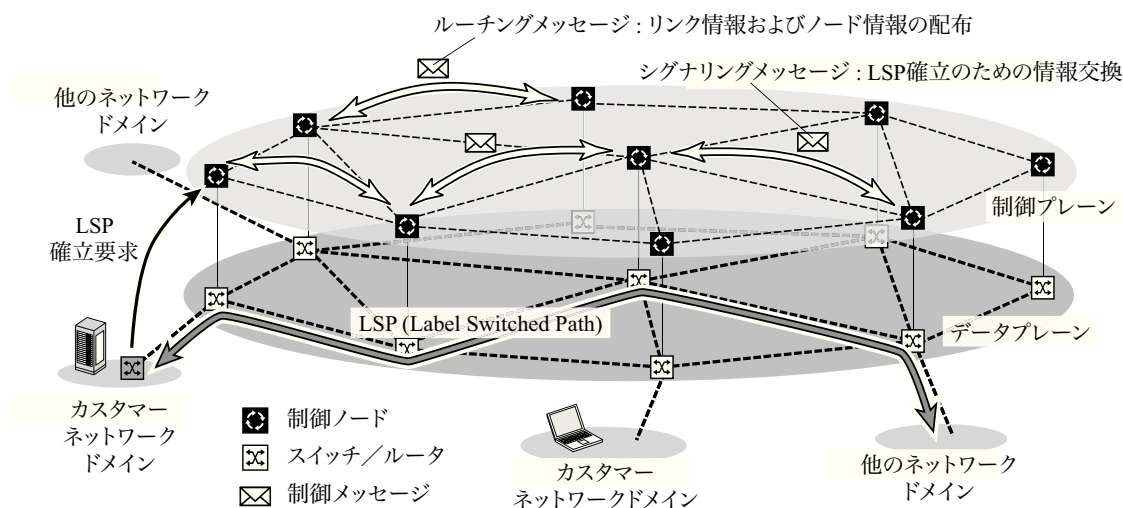


図 4.1: 従来の GMPLS における LSP 確立手法の概要

がデータプレーンである。前者ではルーチングメッセージやシグナリングメッセージの交換，パス計算など LSP 制御を行う。そして，制御プレーンにて LSP 確立が行われた後，全ての LSP はデータプレーン上に展開される。各 LSP は GMPLS の一般ラベル (Generalized Label) と呼ばれる識別子によって区別される。この識別子は，終点ノードアドレスなどのパケットヘッダや，波長，タイムスロットなどが含まれる。

従来の GMPLS における LSP 確立手法を図 4.2 に示す。LSP 確立の最初のステップは，図 4.2(a) に示される制御プレーンにおけるリソース情報の配布である。ネットワークに変更が発生すると，ルーチングプロトコルは各ノードや隣接ノードとのリンクに関するリソース情報をネットワーク上の全ノードへと配布する。この情報はリンク帯域だけでなく，サポートされるメトリック，パケット損失率やエネルギー消費量など，サービス要求を満たす為のノードやリンクの状態および特性が含まれる。ルーチングプロトコルはこれらの情報をフラッディングする際に，重複したメッセージでネットワークを溢れさせないように，効率的にフラッディングを行う。しか

しながら、ネットワーク上の全てのノードに情報を伝える為のメッセージ数は決して少なくはない。

2つ目のステップは図4.2(b)に記される経路計算である。この経路計算のプロセスは始点ノード、すなわち LSP 確立要求を受け取ったノードが行う。一つ前のステップにて、全てのメトリック情報はルーティングプロトコルによって集められ、各ノードのデータベースへと記憶される。そして、確立しようとしている LSP が要求するいくつかのメトリックを使用して、経路計算が行われる。例えば、確立要求が「遅延最小経路」を含む場合、パス計算エンジンは遅延を目的コストとし、最小コストである経路を導出しなければならない。パス計算は CSPF アルゴリズムを用いて行われる。このアルゴリズムは、制約された条件において、指定されたコストを最小化する最短経路を導出可能である。目的コストや制約条件はいくつかのメトリックを用いて表される。しかしながら、GMPLS におけるルーティングプロトコルでは、次の LSP 確立に必要なかどうかに関わらず、サポートされる全てのメトリックが各ノードによって記憶され、最新状態に保たなければならない。

最後のステップは図4.2(c)に表されたシグナリングである。始点ノードはシグナリングメッセージを計算によって求めた経路上の全てのノードへと送信し、LSP を確立する。このメッセージは要求するメトリックなどの情報を含む。LSP 確立のシグナリングメッセージを受信したノードは、自らのスイッチング状態を変更するために関係する情報を取り出し、残りを次のノードへ転送する。メッセージが終点ノードへと到達すると、返信メッセージとして始点ノード宛に返送される。多くの実装では、この返信メッセージが通過する際にスイッチング状態が変更され、データプレーン上にパスが確立される。

この LSP 確立手法の長所は、始点ノードが LSP 確立要求の受け取りから、経路の決定、経路上のノードの制御の全てを担当する点である。これにより、リソー

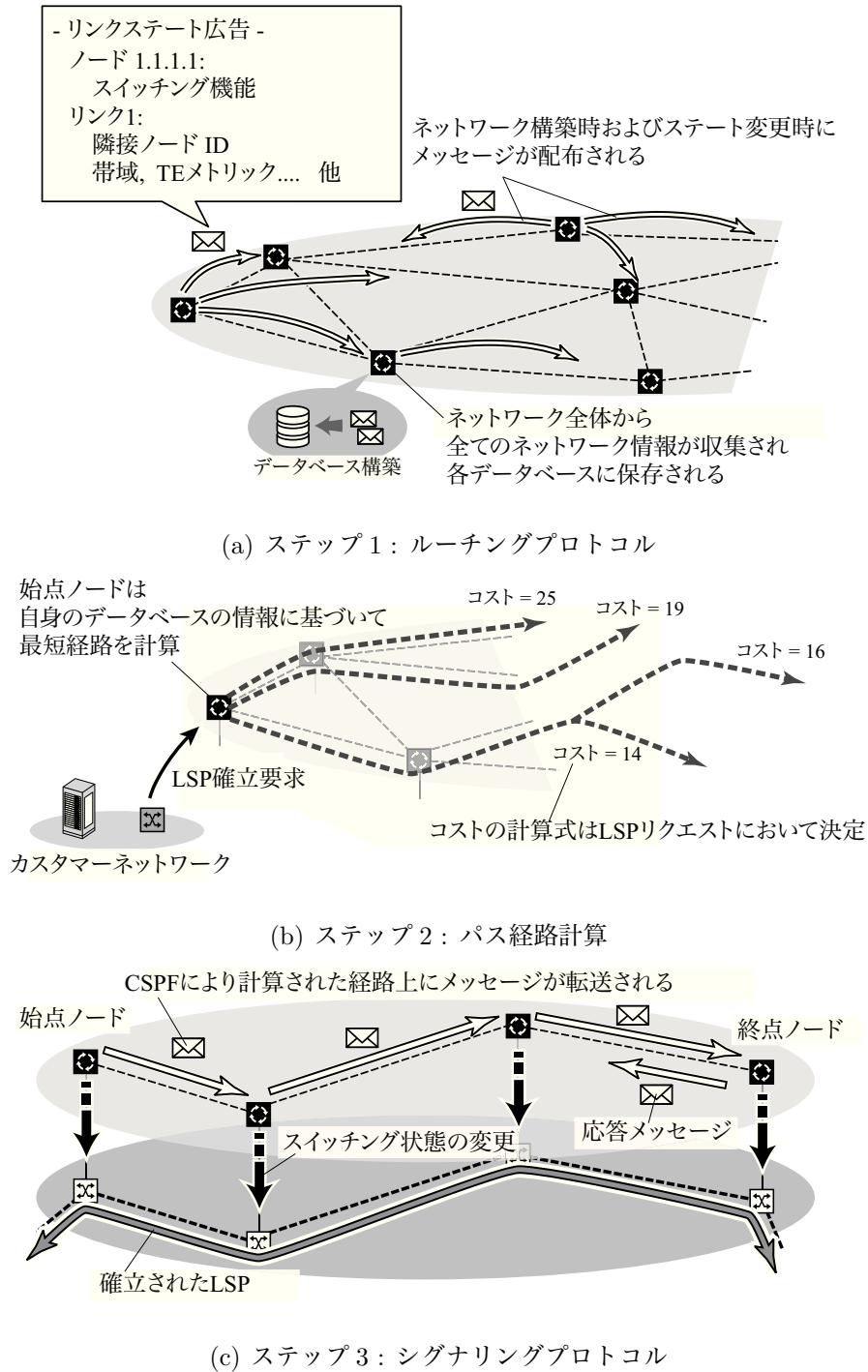


図 4.2: 従来手法における LSP 確立のための 3つのステップ

ス情報を集中管理する必要がない。すなわち、ネットワーク管理者は LSP 管理負荷の局所的な集中を避けることができる。これはネットワーク管理の面では非常に大きな長所となる。

その一方で、この手法は頻繁なリソース情報の広告を必要とする。大量のメッセージがネットワーク上で交換され、これら进行处理するために各ノードのプロセッサのリソースも消費する。もしもメトリック情報が多大だった場合や、ネットワークの変更が頻繁だった場合に、この点は致命的となる可能性がある。

## 4.4 提案手法

### 4.4.1 提案手法の概要

提案手法では、リソース情報をネットワーク上の全ノードへと配布しない。これにより、ルーティングプロセスやリソース情報のためのデータベース、そしてパス計算エンジン無しに LSP 確立が可能となる。そして、FB-RSVP-TE はシグナリングメッセージのフラッディングにより最も適切な経路を探索する。フラッディングでは、各メッセージは通過した経路を記録しながら転送が行われる。通過ノードは、受信するいくつかのメッセージの中から、最も最短経路をたどって来たメッセージだけを選択し、そのノード自身の情報を追記して次へと転送する。また、LSP 要求に対して帯域が不十分なリンクなど、既に条件を満たさない事が明確である場合は転送を行わない。結果的に、最短経路を通過したメッセージが終点ノードへと到着し、CSPF 計算がネットワーク全体で行われたことになる。そして、終点ノードが返信するメッセージが、その経路を通過しながら始点ノードへと戻る時に、LSP が確立される。

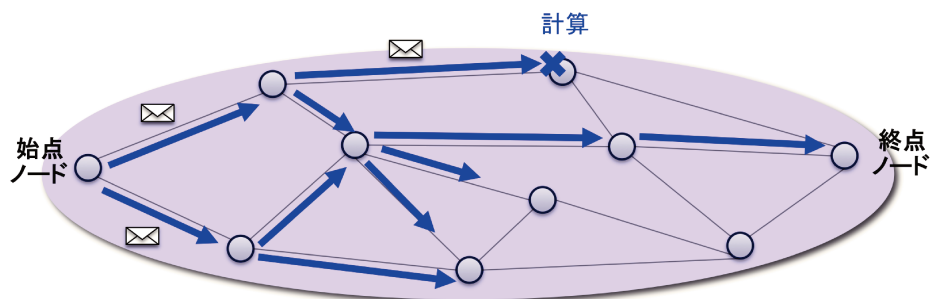


図 4.3: 提案する LSP 確立手法 (FB-RSVP-TE) の概要

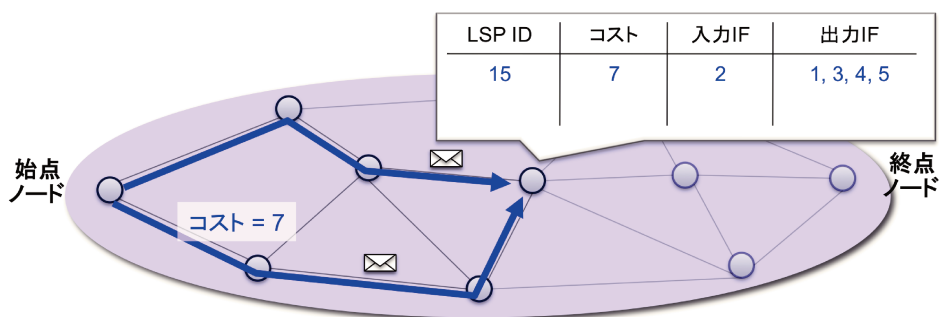
#### 4.4.2 フラッディングに基づくシグナリング手法

FB-RSVP-TE におけるシグナリングの手順を図 4.3 に示す。LSP の確立が要求されると、始点ノードによってフラッディングのためのシグナリングメッセージが作成される。これらのシグナリングメッセージに含まれる情報は、本来の RSVP-TE の PATH メッセージに含まれていたパスに関する情報に加え、新たに定義した経路コスト (Route Cost) フィールドを含む。メッセージが作成された際、このフィールドの値はゼロにセットされ、次のノードへと転送される。そしてメッセージを通過ノードが受信するたびに、各ノードは通過したリンク分のコストを計算し、フィールドの値に加える。このコストは、確立しようとしているパスに関するユーザーやプロバイダの要求を満たした統合メトリックとして計算される。これにより、このフィールドの値は各メッセージが通過した経路の合計メトリックコストを表す。この値によって、各経路の比較が可能となる。各ノードでは、最小コストの経路を通過したメッセージだけが次へと転送される。しかし、最小メトリック探索である以上、このコンセプトは実ネットワークで実現する場合に少し難しさを残す。とりわけ、メッセージのループは絶対に避けなければならない。これらの問題を解決する為に、提

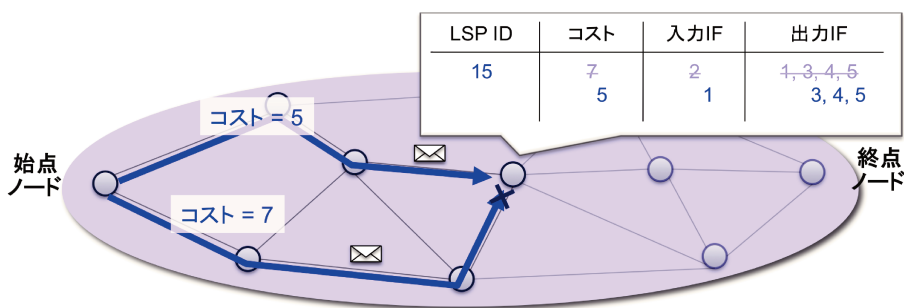


案手法では鍵となる次の2つのメカニズムを含む。一つはシグナリング転送テーブルと呼ばれ、ループを回避しながら最短経路を導出するものである。もう一つはメッセージ追撃と呼ばれ、この最短経路探索の性能を向上させるものである。

### 4.4.3 シグナリング転送テーブル



(a) LSP 確立 (始点から終点へのフラッディング)



(b) LSP 応答 (終点から始点への転送)

図 4.4: シグナリング転送テーブルによるフラッディングおよび経路選択

図 4.4 のように、各ノードはシグナリング転送テーブルを実装する。シグナリング転送テーブルの目的は、最小コスト経路の導出、LSP 確立および応答メッセージ

の転送，メッセージループの回避である。メッセージによってテーブルが更新されると同時にメッセージの転送に使用されるという点で，イーサネットにおける MAC 転送テーブルと非常に似ている。しかし，テーブルの各エントリのキーはネットワーク上で固有なシグナリング ID であって，MAC アドレスではない。これは，各エントリの有効範囲が，1つの確立中の LSP であるためである。

各エントリの最初のフィールドは入力インタフェース，すなわちメッセージの受信ポートである。このフィールドは，エントリが更新された時点での，始点ノードへの最短経路上の方向を示している。二つ目のフィールドは，経路コストである。最短経路の導出は，このフィールドの値との比較によって行われる。すなわち，このフィールド値よりも小さな経路コスト値を持つメッセージが，エントリを更新する。エントリ更新時には，入力インタフェースも更新することにより，入力インタフェースフィールドは常に始点ノードの最短経路方向を示す。一方，大きな経路コストであるために，エントリを更新できなかったメッセージは，破棄される。

エントリを更新したメッセージは，次の隣接ノードへと転送される。このメッセージが送信されるべき隣接ノードの決定にも，シグナリング転送テーブルが利用される。各エントリの最後のフィールドとして用意された出力インタフェースリストが，転送されるべき次の隣接ノードのリストを管理する。確立中の LSP の最初のメッセージの受信した際に，このフィールドには，リンク帯域等の LSP 要求の条件に合致する全ての隣接ノードリストが書き込まれる。そして，メッセージが到着する度に，入力インタフェースの値が，この出力インタフェースリストから取り除かれる。

経路コストは各ノードで常に増加することとなり，また過去に通過したメッセージより大きな経路コストを持つメッセージは破棄されるこのメカニズムにより，メッセージループは確実に回避される。

#### 4.4.4 メッセージ追撃メカニズム

ここまでに説明した通り、各ノードは受信したメッセージの経路コストと、シグナリング転送テーブルの経路コストフィールドの値を比較する。これは始点ノードで経路計算を行う従来のメカニズムとは完全に異なる。

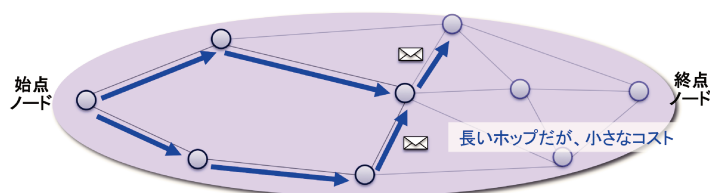


図 4.5: メッセージ追撃メカニズム

図 4.5 にフラッディングベースの提案アルゴリズムの挙動を示す。LSP 確立が要求されると、メッセージのフラッディングを開始し、いくつかのメッセージが始点ノードから隣接ノードへと送信される。これらのメッセージは、次のノードへと接続されたリンクを通過して送信され、受信された後に一定の時間をかけて処理される。ここで、始点ノード以外の各通過ノードでは、全ての入力インタフェースからメッセージを同時に受信する可能性を有するが、これらは同時に処理されるものとする。メッセージが実際にリンク上を通過する時間は、メッセージの処理にかかる時間よりも非常に短いため、結果としてメッセージは図のように 1 ホップずつ進むようにしてネットワーク上に及んで行くと予想される。ここで、もしも長いホップで経路コストが小さい経路が存在した場合、その経路を通過するメッセージは最短ホップを通過したメッセージよりも遅れて到着することになる。時間差のために、この両者のどちらかのみを選択して次に送信することは非常に難しい。そこで、提案では二つ目のメカニズムを導入する。このメカニズムでは、後からやってきた小

小さな経路コストのメッセージは、シグナリング転送テーブルを更新した後に、先行のメッセージを追いかける形で転送される。結果として、いくらかのメッセージの増加を伴いながら、コストの小さなメッセージは終点ノードへと最短経路を伝えるべく到着する。終点ノードではこれらのメッセージのために、経路の最終決定のタイマを作動させる。このタイマの設定時間は、本提案方式における一つの新しいパラメータであり、EWT (Egress Waiting Time : 終点ノード待機時間) と呼ぶこととする。EWT の長さは最短経路の正確さとのトレードオフとなる。短すぎる EWT では実の最短経路を選べない可能性が高まるが、EWT は LSP のセットアップ時間を直接長くする原因でもある。

## 4.5 性能評価

ここまで、従来方式および提案方式における LSP の確立手法について述べた。従来方式ではメトリック更新の度にフラッディングを行う一方で、提案方式では LSP 確立時にのみフラッディングを行う。即ち、メトリック更新が LSP 確立よりも頻繁に行われる環境にて、メッセージ数の削減が予想される。また確立手法の違いによって、LSP 確立のブロック率や、結果として選択された経路、そして LSP セットアップに要した時間についても同時に評価した。本研究では FB-RSVP-TE および OSPF-TE のそれぞれをコンピュータ上でシミュレーションし、各ノードで処理したメッセージ数、LSP 確立のブロック率、選択した経路のスコア、そして LSP セットアップ時間の 4 点について評価を行った。また、始点ノードにおけるセットアップタイマなどその他のパラメータについても考察を行った。シミュレーションプログラムはイベントドリブン型で C++ によって実装された。シミュレーションでは、Waxman[13] モデルを用いて 1000km 四方の仮想的な空間に 100 ノードで 200 リンクのネットワー

クが構築された。各リンクは100Mbpsの帯域を持ち、メッセージの伝達時間は距離を光速で割る事によって導出した。

LSP 確立要求はポワソン到着とし、LSP 要求の平均到着率は1秒につき1とした。ネットワーク負荷は、ネットワーク内の平均LSP数によって定義した。これは、ネットワーク保持時間に要求の平均到着率をかけたものに等しい。各LSP要求が必要とする帯域は指数分布によって与えられ、平均は10Mbpsとした。LSP 確立に関する性能評価を行うため、静的値を統合メトリックとして静的に与えた。提案における経路コストはこの統合メトリックの値を足し合わせることで求められると仮定する。また、統合メトリックとして与えられた静的値は、全てのリンクで等しい一定の値、リンクの距離に基づいた値、そしてランダムに与えられた値の3タイプである。一定の値を与えた場合では、結果として最小ホップ経路が導出され、またリンク距離に基づいた値が使用された場合、経路長が最短となる経路が導出される事が予想される。光ネットワークにおける光の減衰や非線形効果は、このリンク距離に基づいたメトリックに似た挙動を示すと考えられる。またネットワークに関連しないメトリックでは、ランダムに与えられた統合メトリックでのシミュレーションが効果的である。一方、従来手法において、OSPF-TEの更新メッセージはLSP 確立および切断のたびに配布されるものとした。すなわち、メトリックは帯域などのネットワークに関するものであると仮定した。さらに、4.5.1節においてネットワーク非関連である場合の結果を分析した。提案手法におけるEWTの値は1秒、5秒、20秒のそれぞれで結果を求めた。また、全てのシグナリング中の応答メッセージを受け取れるよう、始点ノードにおけるセットアップタイムは説明の無い限り30秒とした。セットアップタイムに関する考察は第4.5.4節にて行う。

表 4.1: メッセージ数を左右する主な要因

	従来手法	提案手法
ルーティングプロトコル	OSPF-TE	N/A
シグナリングプロトコル	RSVP-TE	FB-RSVP-TE
フラッディングを行うプロトコル	OSPF-TE	FB-RSVP-TE
フラッディングが行われるタイミング	メトリックの更新	LSP 確立要求
メッセージ数の近似	ノード数	ノード数

#### 4.5.1 フラッディングメッセージ数

表 4.1 に提案および従来の両手法におけるフラッディングメッセージの主な要因を示した。本節では各プロトコルによって生成されたメッセージの数について評価する。従来手法では二つのプロトコルを実行しており、合計メッセージ数を  $N_{conv}$  とした場合、以下によって表される。

$$N_{conv} = R_r \cdot A_{rsvp} + R_u \cdot A_{ospf} \quad (4.1)$$

式中において、 $R_r$  は要求の平均到着率すなわち単位時間あたりに発生した LSP 確立要求の数、 $R_u$  はメトリック更新頻度すなわち単位時間あたりに更新されたメトリックの数である。 $A_{rsvp}$  および  $A_{ospf}$  は、一階の LSP 確立においてそれぞれ生成されるメッセージ数を示す。メトリックがネットワークに関連したものと仮定した場合、メトリック更新が LSP 確立のみによって発生し、LSP のホップ数を  $H$  とすることによって  $R_u$  は  $2H \cdot R_r$  に等しい。なぜならば、LSP 上の全てのノードが、LSP 確立によるリソースの予約時と、LSP 切断によるリソースの解放時の両方で、メトリック情報の更新を配布する必要があるためである。よって、式 4.1 は次のように

なる。

$$N_{conv} = R_r \cdot (A_{rsvp} + 2H \cdot A_{ospf}) \quad (4.2)$$

一方で、提案方式はルーチングプロトコルを使用しない。そのため  $N_{prop}$  は、 $R_u$  に依存せず、次のようになる。

$$N_{prop} = R_r \cdot A_{fbrsvp} \quad (4.3)$$

本研究では、これらの数式から両手法でのメッセージ数の評価を行った。

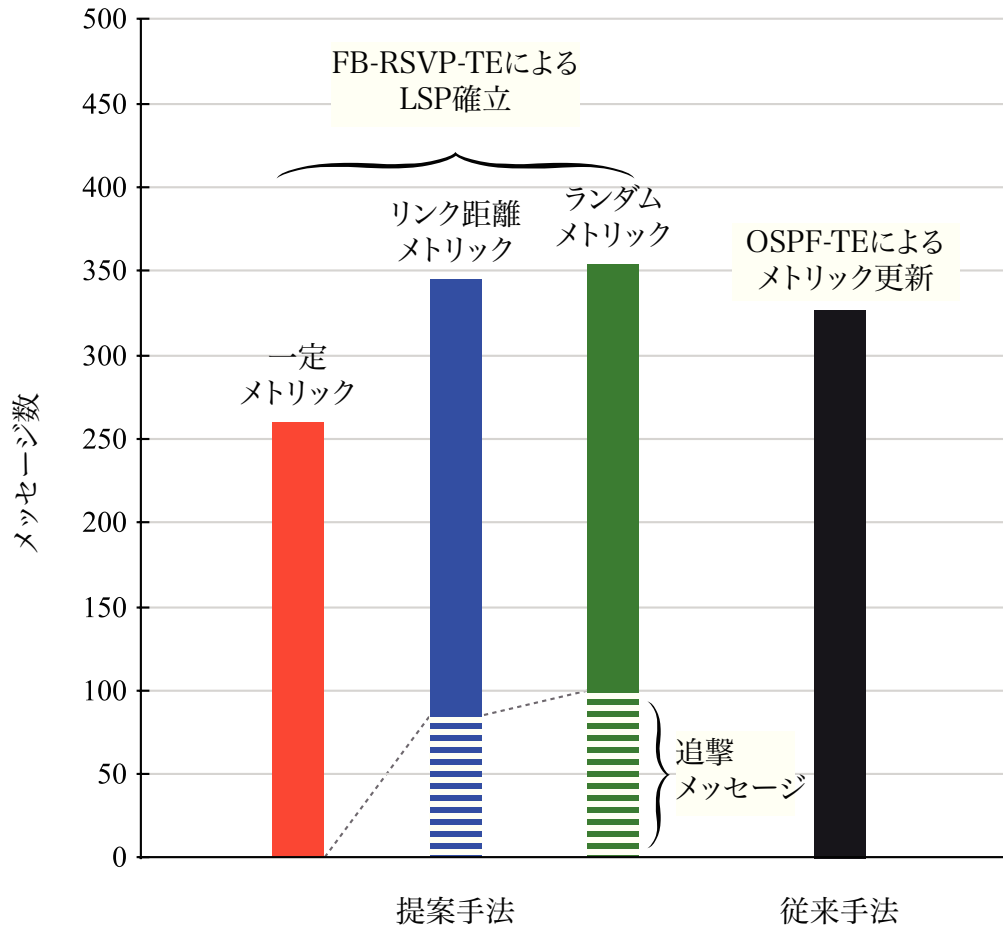


図 4.6: LSP 確立およびメトリック更新毎のフラッディングメッセージ数

図4.6に、シミュレーションによって得られた、ネットワーク上で送信されるフラッディングメッセージ数の違いを示す。提案方式では1回のLSP確立で交換されたメッセージ数であり、従来方式では1回のメトリック更新で交換されたメッセージ数である。それぞれ、 $A_{fbrsup}$  および  $A_{ospf}$  に等しい。図中では、左の3つのバーが提案手法を示す。これらは、統合メトリックとして一定の値、リンクの距離に基づいた値、そしてランダムに与えた場合でのシミュレーション結果である。一方、最も右のバーはメトリック更新における OSPF-TE のフラッディングメッセージである。

提案手法では、一定値のメトリックを与えた場合に最も少ないフラッディングメッセージとなった。これは、遅れて到着する全てのメッセージは長いホップ数を通過したメッセージであり、大きな経路コストとなるために常に破棄され、追撃メッセージが発生しなかった事に依る。その他の二つのメトリック条件では、図のようにいくらかの追撃メッセージが発生する。一定値のメトリックとは異なり、長いホップを通過して遅れて到着するメッセージが小さな経路コストとなる可能性もあり、その場合は追撃メッセージによってメッセージ数が増加する。距離に基づいたメトリックの方が、ランダムに与えたメトリックよりもメッセージ数が少ない傾向にあるのは、短い経路コスト、すなわち短い合計リンク距離は、少ないホップ数によって得られる確率が高い為であると考えられる。一方の OSPF-TE に関しては、OSPF-TE が隣接ノード間で更新による情報の整合性を取り合うプロトコルである以上、一つの更新に対してリンク数と同じだけのメッセージは送信されることとなる。しかし、同じリンク上で同じ方向に2度もメッセージを送ることはない。リンクが双方向であることを考えると、最大でもリンク数の2倍が条件となる。実際は、隣接するノードが互いに同時にメッセージを送信した場合にメッセージ数が増加し、結果として図のようになる。

リンク数 200 と比べ、一定メトリック条件では FB-RSVP-TE はリンク数の 25



パーセント多くメッセージを送信した。そしてランダムメトリックではこれが更に 30 パーセント増加する。

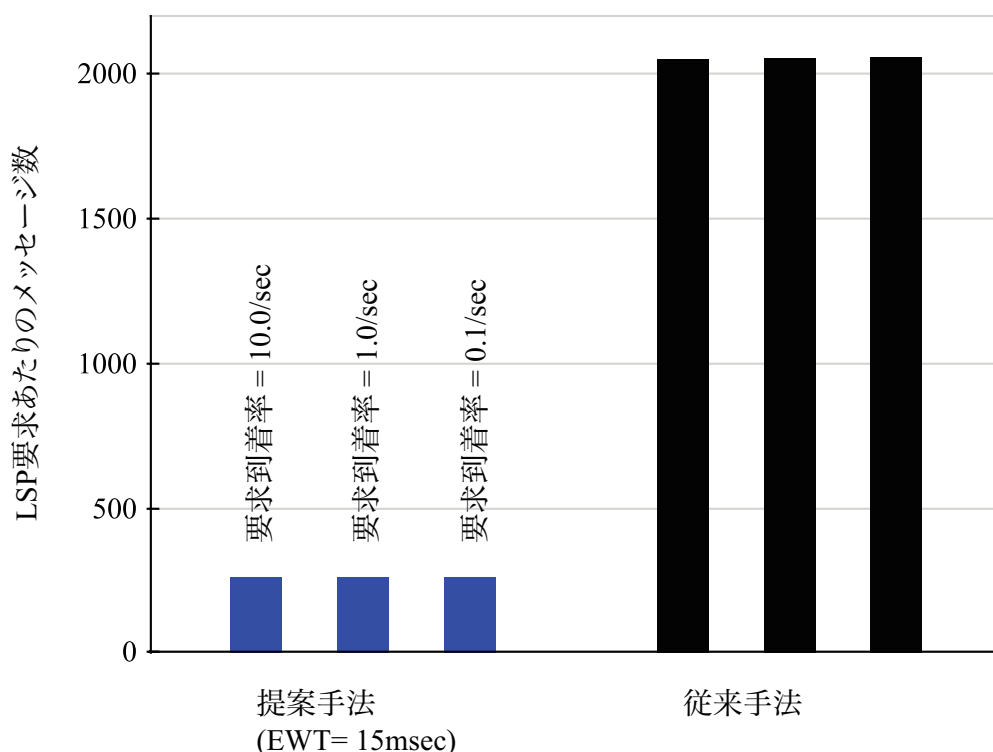


図 4.7: LSP 要求あたりのメッセージ数

このデータを基に、両手法の合計メッセージ数を比較した。まず、メトリックがネットワークに関連したものである場合、提案手法と従来手法のメッセージ数の違いは次の式として求められる。

$$\frac{N_{prop}}{N_{conv}} = \frac{A_{fbrsvp}}{A_{rsvp} + 2H \cdot A_{ospf}} \quad (4.4)$$

ここで、300 個近くのメッセージを生成する OSPF-TE に対し、RSVP-TE はフラッディングを行わず経路上だけに送信される。よって  $A_{rsvp}$  は  $A_{ospf}$  に対して非常に小さい。また、先ほどの結果より、 $A_{ospf}$  はほぼ  $A_{fbrsvp}$  に等しい。これらの近似によっ

て先ほどの数式は次のようになる。

$$\frac{N_{prop}}{N_{conv}} \simeq \frac{1}{2H} \quad (4.5)$$

つまり、提案手法では従来手法に比べて  $1/2H$  までメッセージ数を削減できる。図 4.7 では実験結果における、従来手法と提案手法でのメッセージ数を示す。この図におけるそれぞれのバーは、 $N_{prop}/R_u$  および  $N_{conv}/R_u$  である。また従来手法と提案手法の比  $N_{conv}/N_{prop}$  は 7.89 である。この値から求めると、平均ホップ数  $H$  が 3.869 となる。このように、ネットワークに関連したメトリックであっても、提案手法はメッセージ数を  $1/2H$  に削減する。

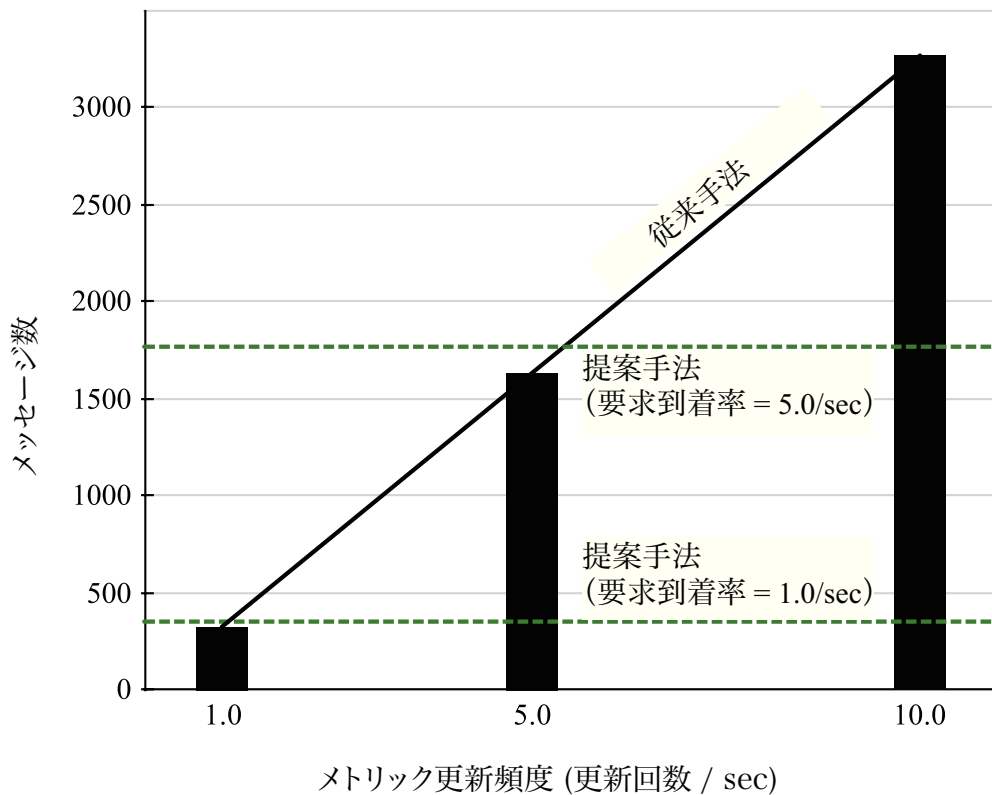


図 4.8: メトリック更新の度に見込まれるメッセージ発生数

一方で、メトリックがネットワークと関連のない場合、LSP 要求とメトリック更新はそれぞれ独立したタイミングで行われる。  $N_{prop}$  と  $N_{conv}$  の比は、次の近似式によって与えられる。

$$\frac{N_{prop}}{N_{conv}} \simeq \frac{A_{fbrsup} \cdot R_r}{A_{ospf} \cdot R_u} \quad (4.6)$$

この数式からメッセージ数の見積もりを示したのが図 4.8 である。要求の到着率が 1.0 および 5.0 だった場合における提案手法によって生成されたメッセージ数が示される。これらの到着率は、100 ノードのネットワークへと、10000 の LSPs が 165 分間もしくは 30 分間に確立された頻度に等しい。これは WDM における頻繁な光パスの確立を想定している。先に述べた通り、  $A_{ospf}$  はほぼ  $A_{fbrsup}$  に等しいため、提案手法と従来手法のメッセージ数の違いは、  $R_u$  と  $R_r$  のどちらが大きいかにゆだねられる。  $R_u$  よりも  $R_r$  の方が大きいケースでは、提案手法がより多くのメッセージを生成する。しかしながら、第 4.2 節で紹介したように、今後の多様化したネットワークでは、  $R_r$  よりも  $R_u$  の方が大きい事が見込まれる。一つのメトリックの更新がルーティングメッセージのフラッディングを伴い、結果として  $R_u$  は全てのメトリックの更新頻度の和となる。これにより、サポートするメトリックの種類が増えた場合、直接  $R_u$  を増加させる。また、一つのメトリックが頻繁なメトリック更新を伴う場合もある。例えば、モバイルバックホールネットワークにおいて、各基地局のセル内にある通信待機状態の携帯電話数は、頻繁に変化する可能性がある。この数は最低でも数秒単位で変化するが、実際に電話をかける頻度、すなわち LSP 要求の到着率はそんなに高くない。別の例として、QoS 評価型ネットワークにおけるトラヒック量がメトリックであった場合、トラヒックの変化は 1 秒以下である。これらのネットワークでは、サポートするメトリックが 1 つであっても、  $R_u$  が  $R_r$  よりも大きい。このように、提案手法は将来のネットワークメトリック多様化において、制御メッセージの数を減らすことが分かる。

表 4.2: 確立ブロッキングの分類

ブロッキングの種類	原因	提案手法での検出方法
リソースの重複予約	複数の LSP 確立が 同一リソースを予約	エラーメッセージ
リソース不足	LSP を確立するための リソースが不足	始点ノードのタイムアウト
セットアップ時間の超過	セットアップタイムで 設定した時間が経過	始点ノードのタイムアウト

#### 4.5.2 ブロック率

次にブロック率に関する評価を行った。ブロック率の評価では、LSP の割り当て方すなわちメトリックの種類によって影響を受けないよう、メトリックは一定値として与えた。LSP 確立のブロックの原因は、表 4.2 に示されるように (1) リソースの重複予約 (2) リソース不足 (3) セットアップ時間の超過のいずれかである。一つ目のリソース重複予約は、同時に発生した複数のシグナリングの衝突である。複数のシグナリングメッセージが同じ経路で LSP を確立しようとした場合、先に LSP 応答メッセージが通過した LSP が確立され、もう一方はブロックを受ける。また、従来手法に限って、誤ったメトリック情報に基づいた経路の選択によっても発生する。OSPF-TE のルーチングメッセージによるメトリック情報の更新が、次の LSP 要求に間に合わなかった場合、始点ノードは古い情報に基づいて経路を計算しなければならない。すると、リソースが残っていない誤った経路へとシグナリングを行

う。両方の場合において、これらはブロックしたノードが始点ノードへとエラーメッセージを送信することによって検出される。二つ目のリソース不足は、ネットワーク負荷が高すぎた場合に発生する。始点ノードと終点ノード間に、LSP 要求が必要とする帯域を残した経路が存在しなかった場合と言い換えても良い。これは、従来手法においては経路計算時に検出され、提案手法では始点ノードにおけるタイムアウトによって検出される。三つ目のセットアップ時間の経過は、シグナリングに時間に対してタイムが短かすぎた場合に発生する。これは適切なタイムの設定という管理上の問題となるため、後の第4.5.4節において考察する。本節ではセットアップタイムを非常に大きな値とし、最初の2つの原因による影響を評価した。

図4.9に提案手法および従来手法におけるブロック率を示す。提案手法では、EWTの値をメッセージ処理時間の整数倍となる4つの値で変化させている。このEWTが小さな値であるほど、ブロック率は小さくなる傾向となる。これは、セットアップ時間が短いことによって、他のシグナリングとの競合の可能性が低くなったためと考えられる。従来手法との比較では、従来手法では提案手法のどのプロットよりも低いブロック率を示した。これは先ほどと同様に、EWTを持たない従来手法では、提案手法よりも常にセットアップ時間が短いためと考える。

EWTの短縮は、経路選択の選択肢の幅を狭める事でもあり、経路の悪化を招くと考えられる。これについての考察は後の第4.5.3節にて述べるが、このトレードオフからブロック率に関しては従来手法の方が少しだけ有利となる。しかしこの差は非常に小さく、提案手法は従来手法とほぼ同等のブロック率を得る事が可能であると言える。

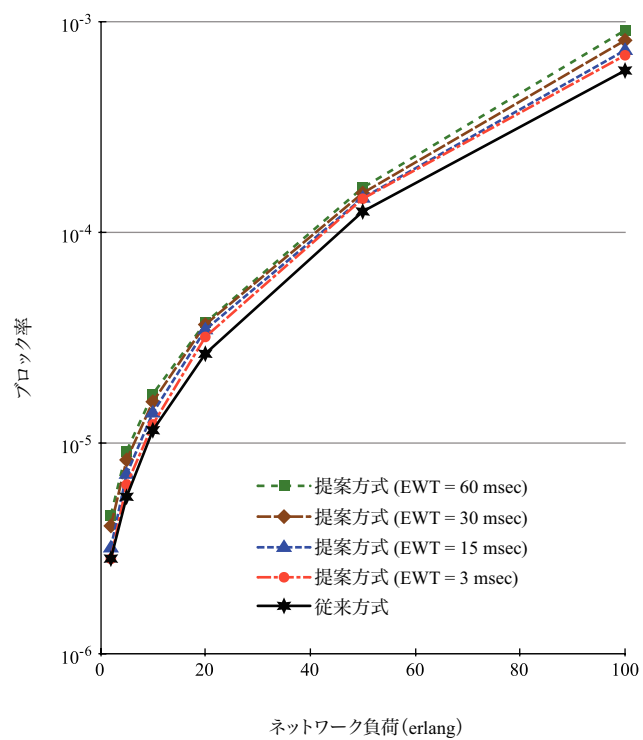


図 4.9: 提案手法および従来手法におけるブロックラ率

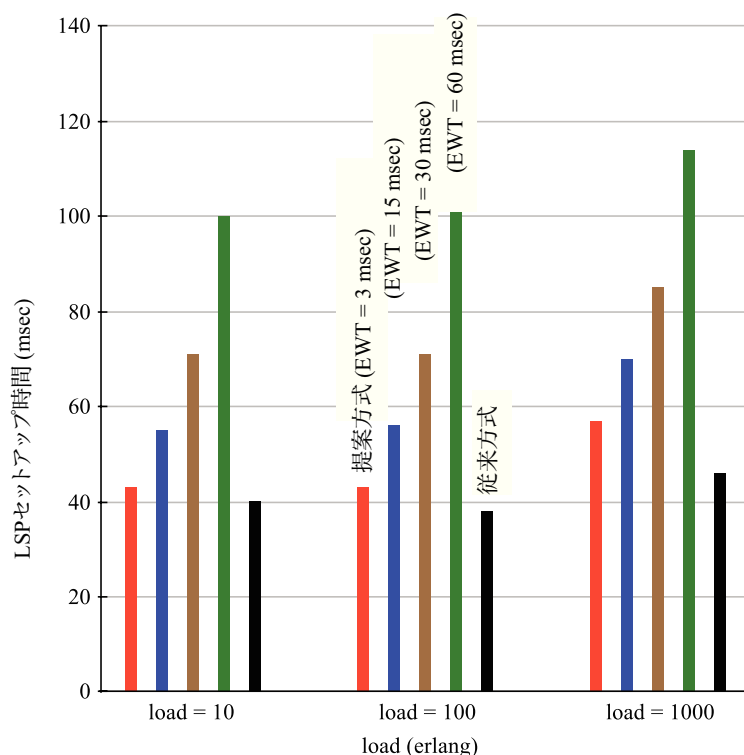


図 4.10: 提案手法と従来手法のセットアップ時間

### 4.5.3 セットアップ時間および経路選択スコア

今度は、LSP セットアップ時間および経路選択スコアについての評価を行った。セットアップ時間は、要求発生時だけでなく、障害発生時における LSP のリストラクション時間にも影響する。経路選択に関しては、ホップ数と最短経路の関連性が低い方が評価の正当性から望ましいため、ランダム値のメトリックを使用した。

図 4.10 にセットアップ時間を示す。EWT を長く設定した場合に LSP セットアップ時間が長くなる事は想定の内である。また、1000 アーラン程度の非常に高い負荷をかけた場合においては、セットアップ時間は増大する。これは、負荷が高いことによって経路が迂回する傾向にあるためと考える。

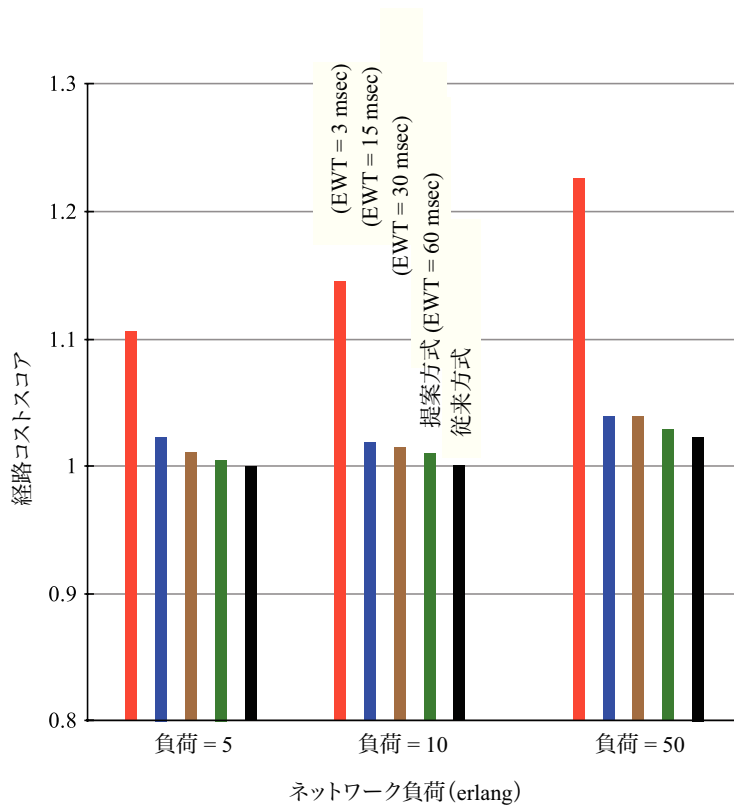


図 4.11: 経路選択スコア

一方の経路選択スコアだが、これは経路の適正さを表す指標として本研究にて定義した値である。これはどれだけ理想の経路から外れたかを表し、実際に確立した全ての LSP の経路コストの合計値を、同じ経路を全て最短経路で確立した場合の経路コストの合計値で割ることによって求められる。すなわち、スコアが1であれば経路は全て最小コスト経路上に確立され、1から離れるほど経路コストの大きな経路が選択されたことになる。経路選択スコアについての結果を図 4.11 に示す。従来手法では負荷が低い場合にスコアはほぼ1となり、提案手法では1より大きな値となる。EWT が最も小さい結果においてのみ、経路が最短経路から大きく外れたことが分かる。また負荷が大きい場合に関しては、いずれの手法においても最短経



路で確立することが困難となるのは明白である。

#### 4.5.4 始点ノードにおけるセットアップタイム

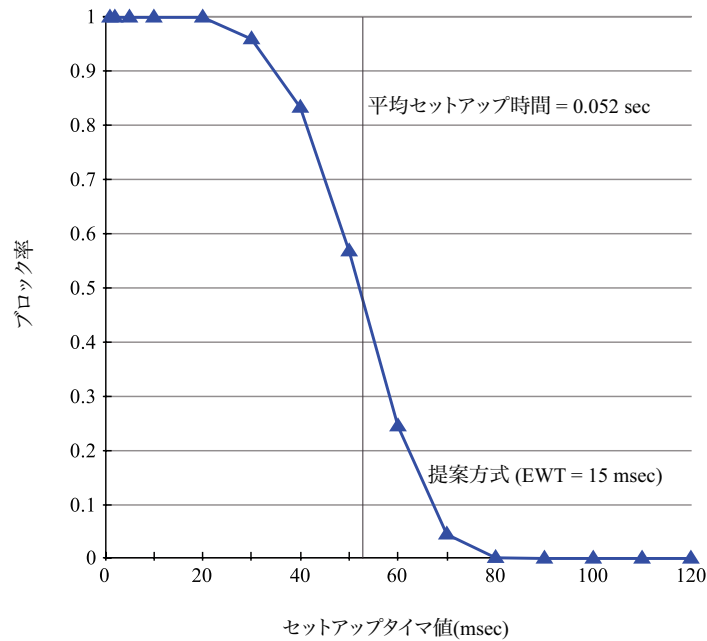


図 4.12: ブロック率とセットアップタイム

最後に、始点ノードにおけるセットアップタイムの値について評価を行った。セットアップタイムは通常シグナリングに何らかのエラーがあった場合、それを検出するために用いられるタイムである。提案手法では、第 4.5.2 節にて説明した通り、リソース不足によるブロックも、このタイムによって検出される。タイムを小さな値に設定した場合、リソース不足によるブロックをより早く検出できる一方で、遠回りした経路によるセットアップ時間の長い LSP 確立をブロックすることとなる。より多くの LSP を受け入れつつ、できる限り早くに検出できるような値が望ましい。

図 4.12 に表される結果では、セットアップタイムの値とブロック率の関係を示す。図のようにセットアップタイムが 100msec 程度において、十分にブロック率が低下したと言える。実際のネットワークを考えた場合、ネットワーク管理者がパス確立の結果に対してどれだけの時間が許容できるかを考えれば、この時間は十分短いと言える。

#### 4.5.5 LSP の張り替えへの影響

ここで、一つの LSP が確立された後に、その LSP を張り替える場合について考える。第 2 章にて議論されたように、GMPLS では提供される固定的な LSP そのものがサービスとなる場合が多い。多くの場合では、要求される度に LSP を確立し、そして不要となった場合に切断する。しかし、通信サービスによってはネットワーク最適化のために LSP の張り替えを行う場合がある。この LSP の張り替えは、ネットワークをより効率的な状態にするために行われる。

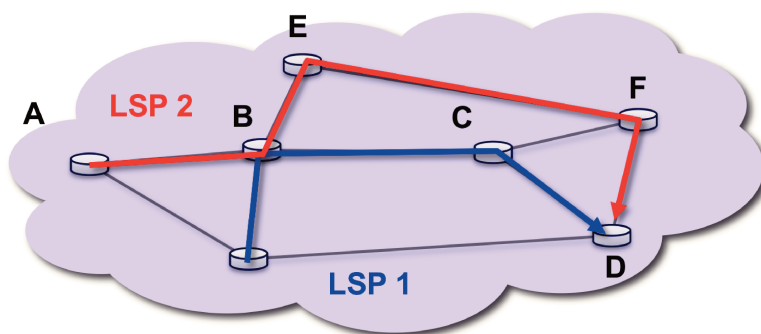


図 4.13: LSP 張り替えの例

提案方式では、LSP を確立するためのシグナリングメッセージがフラッディン

グされる際に、メトリック情報が反映される。しかし、メトリックの更新メッセージを配布しないため、各ノードはネットワーク上の他のリソースに関するリアルタイムなメトリック情報を保持しない。そのため、既に確立した LSP を張り替える余地があるか否か、判断する為の情報を持たない事となる。より具体的な例を図 4.13 に示す。図の例では、最初に確立された LSP1 がリンク B-C におけるリソースを大きく奪ったために、LSP2 は B-C を通過しない遠回りな経路を選択した。ここで、LSP1 が先に切断された場合に、LSP2 が B-C を利用した経路へと切り替えた方が効率が良い事は明らかであろう。提案手法における課題は、LSP1 が切断されたとしても、B-C 間のリソースに関するメトリック情報をノード A が把握することができないという点である。

この課題の対策として考えられる対策の一つは、定期的なりフレッシュを行う事である。一定間隔で同じ始点終点ノード間に同じ条件として LSP のシグナリングを行うことにより、より効率の良い経路があった場合にそれを見つける事が可能となる。その場合、リフレッシュの間隔に比例して、メッセージ数が増大する点を注意しなければならない。4.3 式において表された単位時間あたりのメッセージ数は、リフレッシュ頻度  $R_f$  を用いると次のようになる。

$$N_{prop} = (R_r + R_f) \cdot A_{fbrsvp} \quad (4.7)$$

また、MATE[14] のように複数の LSP を確立し、その LSP 間でトラヒックの流れる割合を変化させるような方式も有効である。この研究では輻輳に関するメトリックのみを扱っており、数本のディスジョイントな LSP と、5 分おきの定期的なプローブパケットによって十分な性能が得られる事を示している。しかし、本研究で扱うような、電力量など他のメトリックに関する影響は、検討の余地があるだろう。

## 4.6 結論

本章では，GMPLS ベースの TE における，メトリック多様化に対する規模拡張性について扱った。アプリケーションのための QoS 保証や，物理レイヤーを考慮した TE の拡張によるマルチメトリック化は，メトリックの頻繁な更新を発生させる。メトリック更新の度にネットワーク全体へとメッセージを配布する従来方式では，頻繁な更新が行われた場合にネットワークがメッセージに逼迫されてしまう。そこで，提案手法である FB-RSVP-TE では，シグナリングメッセージをフラッディングすることにより，ルーチングプロトコルによるメトリック情報の配布を行わずに最短経路の導出を可能とした。この提案手法では，従来手法とほぼ同等のブロック率を保ちながら，メトリック更新の毎にメッセージを発生させない。コンピュータシミュレーションを用いた評価によって，メトリック更新頻度の高いネットワークにおいて，メッセージ数を大幅に削減することを示した。

## 参考文献

- [1] E. Mannie, et al., “Generalized Multi-Protocol Label Switching (GMPLS) Architecture,” IETF Proposed Standard, RFC 3945, October 2004.
- [2] E. Rosen, et al., “Multiprotocol Label Switching Architecture,” IETF Proposed Standard, RFC 3031, January 2001.
- [3] D. Katz, et al., “Traffic Engineering (TE) Extensions to OSPF Version 2,” IETF Proposed Standard, RFC 3630, September 2003.
- [4] D. Awduche, et al., “RSVP-TE: Extensions to RSVP for LSP Tunnels,” IETF Proposed Standard, RFC 3209, December 2001.
- [5] Siamak Azodolmolky, et al., “A Survey on Physical Layer Impairments Aware Routing and Wavelength Assignment Algorithms in Optical Networks,” *Journal on Computer Networks and Telecommunications Networking* , Volume 53, No. 7, Page 926-944, May 2009.
- [6] Y. Lee, et al., “A Framework for the Control of Wavelength Switched Optical Networks (WSO) with Impairments,” IETF Internet draft, draft-ietf-ccamp-wson-impairments-07, April 2011.

- [7] Y. Lee, et al., “Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks,” IETF Internet draft, draft-ietf-ccamp-rwa-info-11, March 2011.
- [8] Luis Velasco, et al., “Failure Aware Diverse Routing: A Novel Algorithm to Improve Availability in ASON/GMPLS Networks,” Transparent Optical Networks 2006 International Conference, Page 195-198, June 2006.
- [9] Haruka Yonezu, et al., “QoS Aware Energy Optimal Network Topology Design and Dynamic Link Power Management,” European Conference and Exhibition on Optical Communication 2010, Tu.3.D.4, September 2010.
- [10] Antonio Cianfrani, et al., “An Energy Saving Routing Algorithm for a Green OSPF Protocol,” INFOCOM IEEE Conference on Computer Communications Workshops 2010, May 2010.
- [11] Daisuke Ishii, et al., “A Novel IP Routing/Signaling Based Service Provisioning Concept for Ubiquitous Grid Networking Environment,” IEEE Globecom 2010 Workshop on Ubiquitous Computing and Networks, No. 4, Page 1808-1812, December 2010.
- [12] D. Johnson, et al., “DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks,” Ad hoc networking (Edited by Charles E. Perkins), Page 139-172, December 2000.
- [13] B. M. Waxman, “Routing of multipoint connections,” IEEE Journal on Selected Areas in Communication, Volume SAC-6, Page 1617-1622, December 1988.

- [14] A. Elwalid, et al., “MATE: MPLS Adaptive Traffic Engineering,” Proceedings of IEEE INFOCOM, Page 1300-09, April 2001.

---

## Chapter 5

---

# 広域イーサネットにおける トラフィックエンジニアリングのための ドメイン連続性に関する規模拡張性

### 5.1 概要

ここまでの第3章および第4章では制御プレーンのTEを扱ってきた。第3章ではIPベースのTEに、また第4章ではGMPLSのTEについて着目した。ここからは、下位のレイヤーと関連したTEに関する研究について扱う。特に、伝送レイヤーの技術であるイーサネットをWANにおいて利用した広域イーサネットに関する規模拡張性の問題に注目した。本章では、広域イーサネットにおいてTEを実現する為の、ドメイン連続性に関する規模拡張性の問題を解決する。広域イーサネットにおいては、パスを確立するVLANのID空間の制限によりネットワーク全体で接続性に制約を持ち、複数ドメインにおける接続確立が困難であった。そこで、ドメイン境界ノードにおいてタグをスワップさせることによりドメイン毎のID空間を独立化を実現する。VLANタグのスワップにともない、GMPLSにおけるシグナリング機能の拡張が必要となる。本章では、マルチドメインによって構成されるネットワークにおけるVLANタグのスワップを含むシグナリング方法につい



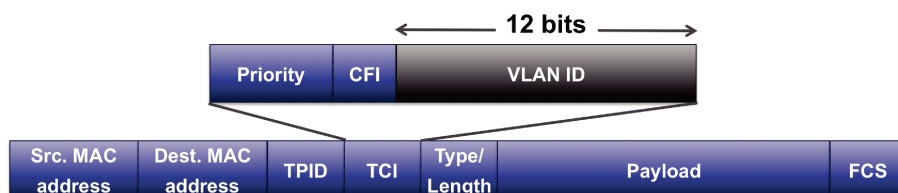


図 5.1: VLAN フレーム

て、ISP におけるリソース秘匿性の制約をふまえた二つのシグナリング手法を提案した。それぞれの手法は、プロトタイプネットワーク上へと実装し、適切に VLAN スイッチを制御する事によって VLAN パスの確立することが実証された。

## 5.2 広域イーサネットにおけるマルチドメイン

### VLAN パス確立

広域イーサネットは、次世代のインターネットにおける伝送レイヤ技術として期待されている技術である。本来 LAN における通信技術として開発されたイーサネットを用いる事により、費用対効果が高くシームレスな通信が実現される。しかし、広域ネットワークは多くの点で LAN と異なっている。例えば信頼性、QoS、管理方法、パスの制御、障害検出や OAM 機能は LAN においてはあまり必要とされないが、広域ネットワークでは必要な不可欠な機能である。これらの広域イーサネットにおける課題や取り組みについては Akira Arutaki ら [1] によってまとめられている。

広域イーサネットの主な利用は、到着するイーサネットフレームを宛先へとそのまま転送する仮想線サービスである。この仮想線は VLAN の技術によって識別され、VLAN パスと呼ばれる。イーサネットスイッチは MAC アドレスだけでなく、VLAN 拡張されたイーサネットフレーム内の VLAN ID に基づいてフレームを転送する。

VLAN ID とは、図 5.1 に表されるように VLAN 拡張された Ethernet フレーム上の ID 値である。各スイッチへと VLAN を適切に設定することにより、物理ネットワーク上に VLAN パスが確立される。VLAN パスを確立および切断する際には、パス上の全てのイーサネットスイッチが適切に設定されなければならない。GMPLS はスイッチの設定を自動的に行い VLAN パスを LSP として確立する制御プレーン技術として適用可能である。すなわち、GMPLS を使用することで、要求に応じた VLAN パスの確立自動化を実現できる。GMPLS による広域ネットワーク上での VLAN パス確立はネットワークを運営する ISP にとって非常に大きな挑戦である。そのために、Internet Engineering Task Force (IETF) が自動 VLAN パス設定プロトコルである GELS(GMPLS Ethernet Label Switching) [3] のフレームワーク確立を開始している。GELS では、先に述べたような広域ネットワークにおいて求められる機能について言及しているが、具体的な識別子リソースである VLAN ID の管理については触れていない。VLAN は LAN における利用を考慮して拡張されたため、ID 空間は非常に小さく、MPLS のラベルと同じように扱う事はできない。そのため、広域イーサネットにおいて、この ID 空間が VLAN パス確立の制約となる可能性もある。特に、複数のドメインを跨いだ VLAN パスを確立する場合に、ID の衝突が課題となる。

本章では、マルチドメイン環境における広域イーサネットでの VLAN ID の管理について注目した。マルチドメインにおける GMPLS 制御広域イーサネットにおいて VLAN タグのスワップ機能を実現し、実際にプロトタイプを実装した。このプロトタイプにおける実験にて、(1)VLAN ID 指定によるパス確立、(2) 任意のスイッチにおける VLAN ID のスワップ、(3) ドメイン毎の VLAN ID の管理の 3 つの機能を実証した。

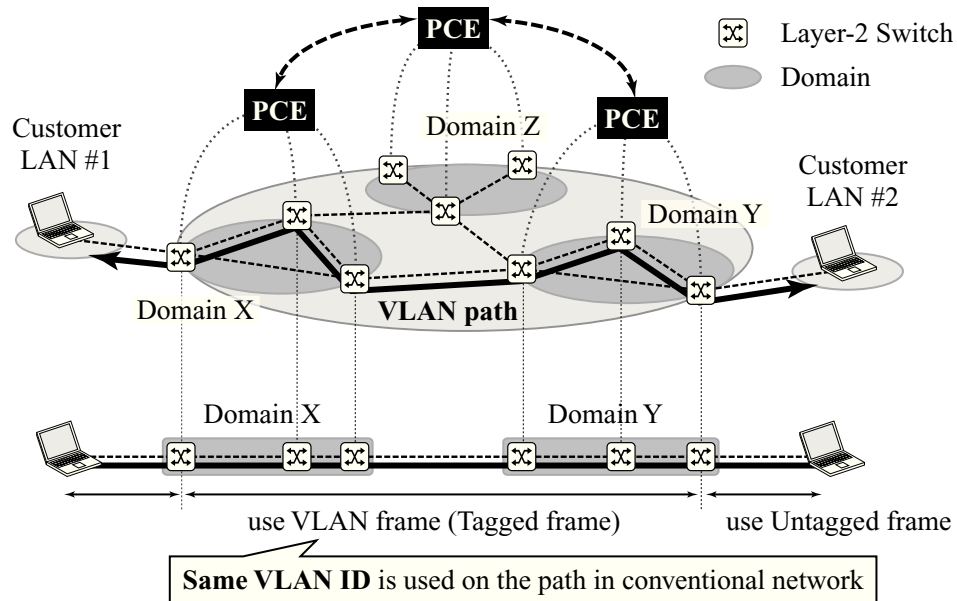


図 5.2: GMPLS 広域イーサネットのアーキテクチャ

### 5.3 広域イーサネットのアーキテクチャ

図 5.2 に本研究にて想定するマルチドメインの広域イーサネットのアーキテクチャを示す。このネットワークは GMPLS によって制御される。図に示される PCE は、各ドメインの全ノードにおけるリンクやスイッチなどの全てのリソース情報を管理する機器であり、図中では各ドメインの上に位置するが、実際にはいずれかのノードに実装されることが多い。PCE が管理するリソース情報には VLAN タグリソースに関する情報も含む。ネットワーク上への新たな VLAN パス確立は、そのパスの始点となるノードへと要求される。図中の例では、カスタマー LAN#1 からカスタマー LAN#2 へのパスがリクエストされている。そして、始点ノードからパス確立の処理が開始され、シグナリングによってパスが確立される。

従来の広域イーサネットは、パス上で全て等しい VLAN ID 値を使用する必要があった。これは、イーサネット標準 [4] において異なる VLAN ID 値を一つの VLAN

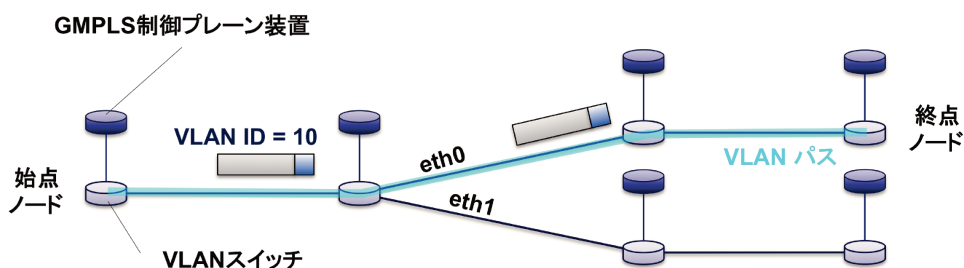


図 5.3: VLAN パス

グループに使用できないという基準に従ったためである。一方で、VLAN ID 空間 12 ビットであり、ID の取りうる範囲は 1 から 4096 に限られている。そして想定されるネットワークサイズは非常に大きい。この VLAN ID の制限は複数のキャリアが運用するマルチドメインの広域ネットワークに対して、非常に厳しい制約となる。

この問題を解決するために、Q-in-Q と呼ばれる VLAN タグをスタックする技術や、MAC-in-MAC と呼ばれるカプセリング技術が開発された。GMPLS はこれらの Q-in-Q および MAC-in-MAC の技術へと適用する事も可能である。しかし、エンドツーエンド（ユーザーからユーザー）における VLAN パスの確立が要求されている場合において、マルチレイヤーの GMPLS シグナリングの実装は技術的に複雑である。この複雑なマルチレイヤーシグナリングの実装を避ける為に、本研究では VLAN タグのスワップを用いたパス確立に挑戦した。ドメイン境界において、VLAN ID をスワップすることにより、ドメイン毎に独立した VLAN ID を割り当てる事が可能となる。これにより、より規模拡張性に優れた VLAN ネットワーク設計が可能となる。

本章では、VLAN のスワップ機能をプロトタイプであるレイヤー 2 スイッチ上に実装した。スワップ可能なプロトタイプのスイッチと、従来の通常のスイッチとの間の相互接続のために、GMPLS RSVP-TE におけるラベルの使い方を拡張した。

## 5.4 VLAN パスの確立

本章では、ドメイン境界でのタグスワップを用いた VLAN パス確立を行うための、各ノードの制御について議論する。最初に、ドメイン境界でのタグスワップを行わずに各一つのドメイン内での VLAN パス確立について述べる。次に、マルチドメイン間においてドメインタグスワップ技術を用いた 2 つの VLAN パス確立手法について述べる。

### 5.4.1 単一ドメイン内における VLAN パスの確立

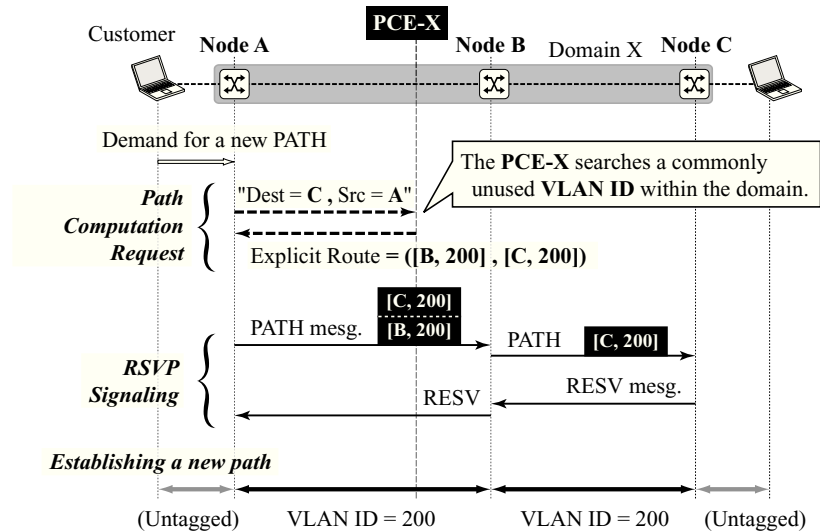


図 5.4: 単一ドメイン内における VLAN パス確立のシーケンス図

図 5.4 へと、単一のドメイン内でパスを確立する際の各ノードの動作について示した。図中では始点ノードであるノード A がパス確立を開始する。パス確立は大きく分けて 2 つのプロセスが必要となる。一つ目はパスの経路を決定するプロセスである。このネットワークでは PCE による集中管理が想定されているので、PCE

へと経路計算リクエストを送信することによって経路は決定される。パスの経路は、経路上の全てのノード ID と VLAN ID の情報として獲得する。二つ目の処理はこれらの情報に基づく RSVP-TE によるシグナリングである。シグナリングでは、PATH メッセージを下流ノードへと送信し、イーサネットスイッチを設定する事によってパスを確立する。このメッセージがパス上の全てのノードを通過した後、全てのスイッチとリンクにおいて VLAN ID が確保される。

交換されるメッセージについてももう少し詳しく注目する。VLAN パスの確立要求は、始点ノードと終点ノード、そして必要な帯域が決まっている。この要求が始点ノード A に到着すると、ノード A はどのドメインのリソースを管理する PCE X へと、終点ノード C のアドレスと帯域についての情報を経路計算リクエストのメッセージとして送信する。PCE-X による終点ノード C までのリソースの使用状況を考慮して経路を導出し、まだドメイン X 内で未使用な VLAN ID を見つけて、ノード A へと詳細な経路情報を返信する。この経路情報は VLAN ID を含み、【B,200】【C,200】のように次のノードのアドレスと、VLAN ID をセットにしたもので表現され、ERO(Explicit Route Object) と呼ばれる。ノード A はこの ERO を含む RSVP-TE の PATH メッセージを作成し、またその先頭に記述されたノード B へと転送する。ノード B やノード C など他のノードはメッセージを受け取るとその ERO に基づいてレイヤー 2 スイッチを適切に設定する。そしてメッセージが終点ノード C へと到着すると、ノード C はこれを RESV メッセージに変更し上流へと返信する。この RESV メッセージを各ノードを通過する際に、実際にスイッチの VLAN 設定が変更され、VLAN パスが確立される。

## 5.4.2 マルチドメインにおける VLAN パス確立

本節では，VLAN パスを確立するための二つの提案方法について，図 5.5 を用いて説明する．マルチドメイン環境では，それぞれのドメインにおける経路の決定とシグナリングのタイミングの違いから二つの手法が提案される．

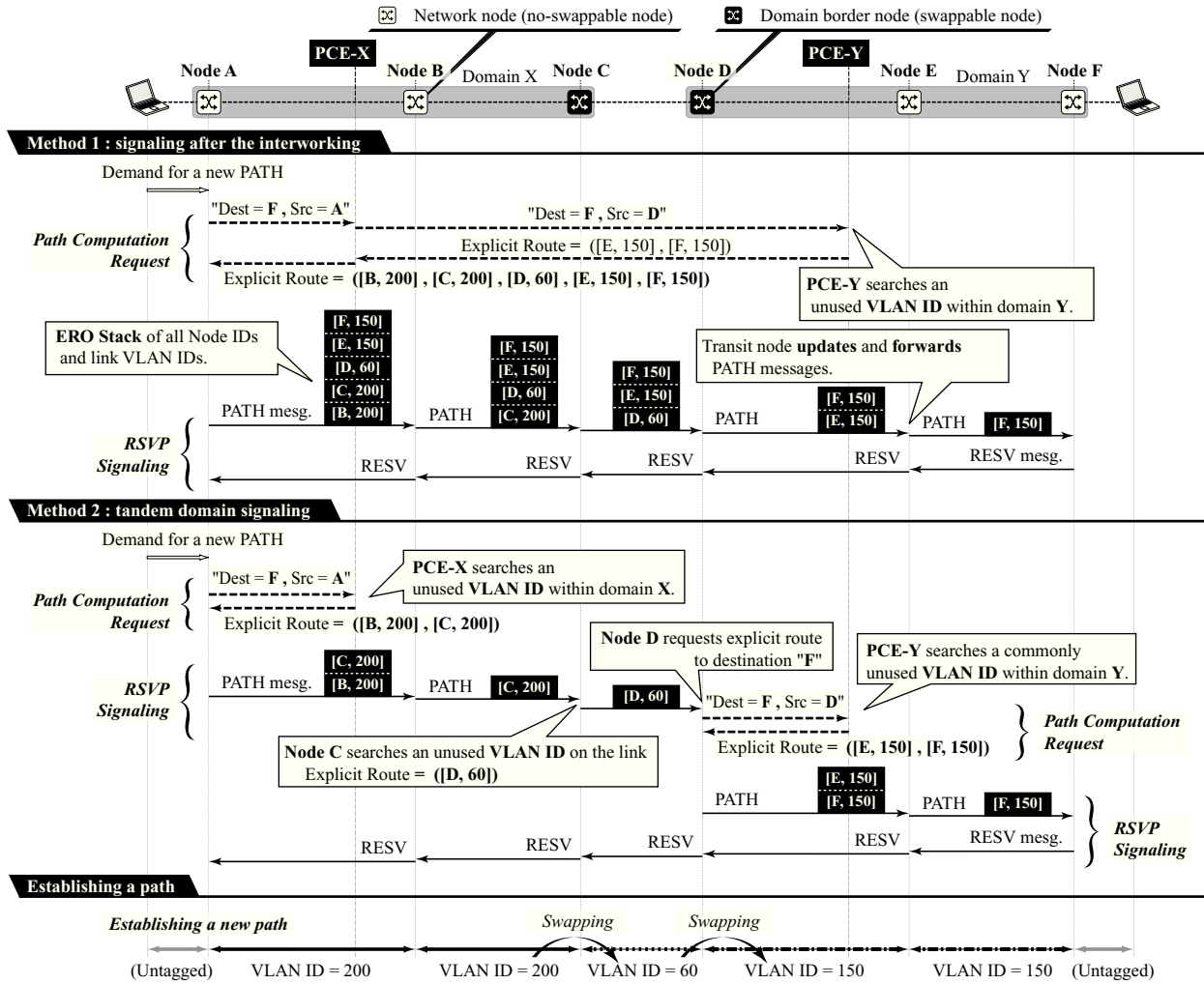


図 5.5: マルチドメインにおけるパス確立のシーケンスダイアグラム



#### 5.4.2.1 方式 (1) : ルーティング後のシグナリング

方式 (1) では、経路計算および RSVP-TE シグナリングはそれぞれのフェーズで行われる。最初に、始点ノード A は、ノード F までの経路のために、パス計算リクエストメッセージを PCE-X へと送信する。ここで、ノード F はドメイン X 内ではないため、複数のドメインをまたいだ経路が必要になる。PCE-X はあらかじめ BGP(Border Gateway Protocol) など他のプロトコルによって、終点ノード F が所属するドメインを把握している。そのため、ノード F が所属するドメイン Y の PCE-Y とメッセージの交換を行いながら経路を決定する。

経路計算は、PCE-X と PCE-Y のそれぞれで行われる。まず始点ノード A は、ドメイン Y へのゲートウェイとなるノード C への経路を計算する。そしてノード C が接続するノード D 以降の経路を PCE-Y へと一任する。PCE-Y へと経路計算リクエストを送信し、ノード D からノード F までの経路を取得する。そして、PCE-X は二つの経路情報を組み合わせ、A から F までの ERO としてノード A へと返信する。

境界ノードにおいてドメインタグスワップ機能をサポートするために、ドメイン毎に異なる VLAN ID が使用可能となる。図中では、VLAN ID = 200 がドメイン X へと割り当てられ、VLAN ID = 150 がドメイン Y へと割り当てられる。そしてドメイン間は VLAN ID = 60 である。パス計算リクエストによって詳細な経路を取得した後、ノード A は PATH メッセージを作成し、下流ノードへと送信する。PATH メッセージの受信したノードは、VLAN ID の情報を獲得し、レイヤー 2 スイッチを制御してパスを確立する。

この方式では、ルーティングとシグナリングは独立して行われる。PATH メッセージを受信した各通過ノードは、メッセージ内の ERO オブジェクトを読み込み、そしてメッセージを次の下流ノードへと転送する。

この方式の欠点は、リソース情報であるドメインの経路情報が、始点ノードほ

か他ドメインのノードに知られてしまうことである。この問題はネットワークを運営する際に機密性の課題となる可能性がある。

#### 5.4.2.2 方式(2)：タンデムドメインシグナリング

2つめの方式はタンデムドメインシグナリングでは、PCEは自身のドメイン内の経路のみを扱う。PATHメッセージが次のドメインに到達するたびに、ドメイン境界ノードはPCEに経路計算リクエストメッセージを送信する。

図5.5において、新たな経路の確立要求を受けた後、始点ノードAは送信元ノードとゲートウェイノードC間の経路計算をPCE-Xに要求する。PCE-XはAからCへの経路を計算し、ドメインX内で未使用のVLAN IDを検索し、経路確立のための明示的な経路情報を作成して返信する。始点ノードAはノードCまでの情報を受け取ると、ここでRSVP-TEシグナリングを開始する。ノードCに到達するPATHメッセージを受信すると、ノードCは次のリンクに対する未使用VLAN IDを検索する。リンクのリソース情報は、リンクの両側のノードによって管理されている。ノードCはメッセージにEROを追加し、ダウンストリームノードDに送信する。ノードDはPCE-Yにリクエストメッセージを送信する。PCE-YはDからFへの経路を計算し、ドメインY内において未使用のVLAN IDを検索、新たな経路を確立するための明示的な経路情報を作成する。ノードDはメッセージにEROを追加し、ダウンストリームノードEにメッセージを送信する。その後、PATHメッセージが出口ノードFに到達し、ノードAにはRESVメッセージが応答されることにより、新たな経路が確立される。

本方式は、PCEが他のPCEと相互接続することなしに経路の計算を実行するため、経路の計算は各ドメイン内で独立的に実行される。また、VLAN ID等の全てのリソース情報はドメイン内で伝達され機密性を保持している。

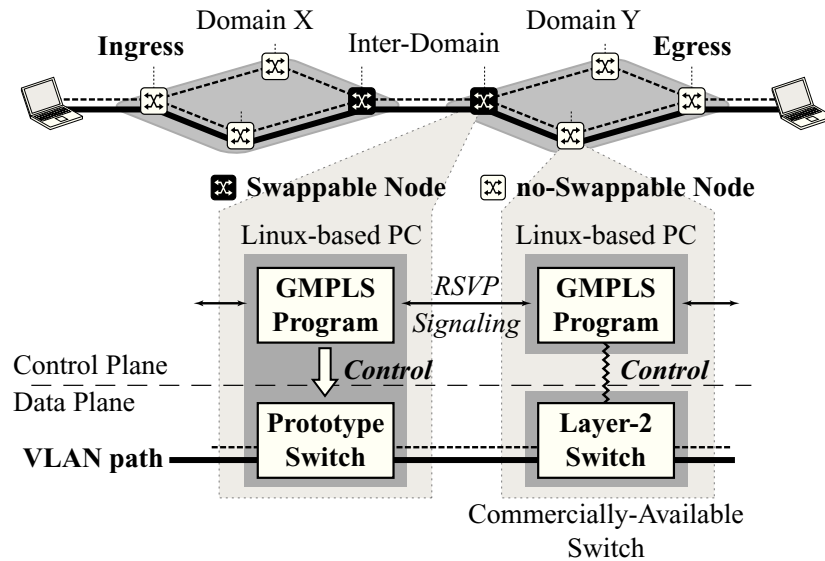


図 5.6: 実験ネットワーク

## 5.5 特性評価

### 5.5.1 実験ネットワーク環境

図 5.6 は、実験ネットワークを示している。本ネットワークは、図 5.2 で示した広域イーサネットのマルチドメインの相互接続を実装している。ネットワーク中には 2 つのドメインが存在し、各ドメインには 4 つのノードが存在する。VLAN タグスワップ機能を搭載したスワップ可能なスイッチをドメインの境界に設置し、他の 6 つのスワップをサポートしないスイッチは商用のイーサネットスイッチで構成される。それぞれの制御は Linux PC 上で動作するプロトタイプの GMPLS プログラムによって行う。適切な PATH, RESV メッセージを受信した場合、全てのスイッチはエンドツーエンドの VLAN パスが正常に確立される。

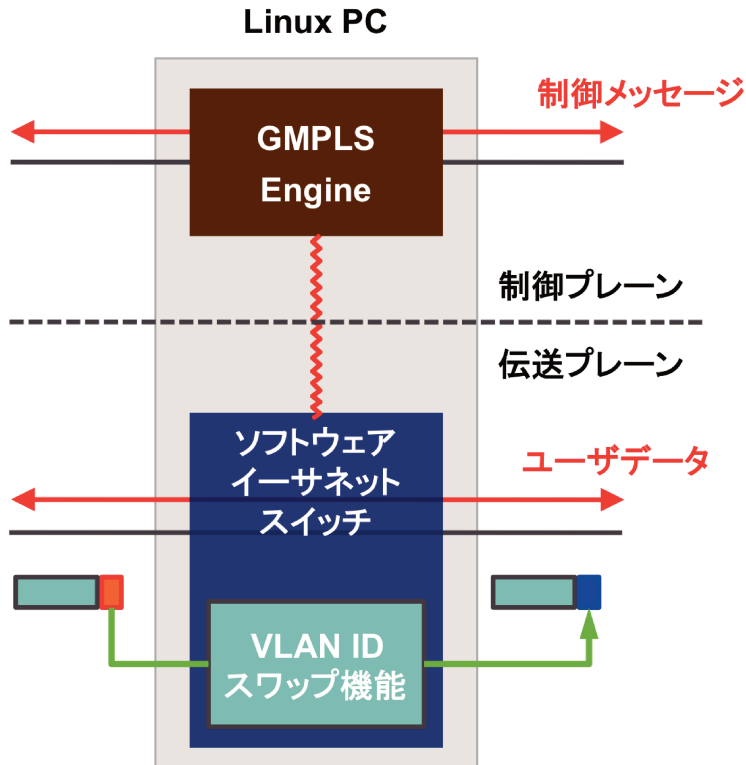


図 5.7: 各ノードの実装

## 5.5.2 タグスワップを実現したマルチドメイン

### VLAN パスシグナリング

実験において、5.4節で述べた2つの方式でVLAN タグスワップパス確立を行った。実装されたRSVP-TEプログラムは、ルーターのIDだけではなくラベル[5]も含むEROを指定するように拡張した。各リンクのVLAN IDはGENERALIZED LABELオブジェクトによって扱われる。このフォーマットは[6]に示されている。

本稿におけるドメインタグスワップ機能を搭載したVLANパス制御メカニズムは国際会議iPOP 2008の展示において展示され、成功を取めた。両シグナリング方

式とも、商用で手に入るイーサネットスイッチとプロトタイプスイッチとの相互接続が可能なことを実証した。

## 5.6 結論

本章では、広域イーサネットにおけるマルチドメイン TE を実現する為の VLAN ID のドメイン連続性に関する規模拡張性問題に注目した。そこで、シグナリング機能を拡張し、ドメインタグスワップによりドメイン毎の ID 空間を独立化を行った。その際、マルチドメイン VLAN パスシグナリングにおける 2 つの方式を提案した。1 つ目は、PCE の相互接続を利用した方式であり、2 つ目は、タンデムドメイン方式である。本方式は、将来的な広域イーサネットに適応可能であることを示した。

## 参考文献

- [1] Akira Arutaki, “Wide Area Ethernet Technology Overview (in Japanese),” IE-ICE book, ISBN:4-88552-211-0, June 2005.
- [2] E. Mannie. (Editor), “Generalized Multi-Protocol Label Switching (GMPLS) Architecture,” IETF RFC 3945, October 2004.
- [3] D.Papadimitriou, et.al, “Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE)”, IETF Proposed Standard, RFC 6003, April 2008.
- [4] IEEE Std 802.1Q, “IEEE Standards for Local and Metropolitan Area Networks Virtual Bridged Local Area Networks,” May 2003.
- [5] L. Berger (Editor), “Generalized Multi-Protocol Label Switching(GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering(RSVP-TE) Extensions,” IETF RFC 3473, January 2003.
- [6] D. Ishii et al, “An Experiment of Controlling Gigabit Wide Area Ethernet by GMPLS supporting Layer-2 Switching Capability”, Proceedings of OFC NFOEC 2008, OThB5, February 2008.

- [7] K. Kikuta, et al., “GMPLS VLAN Path Establishment using Inter-domain VLAN Tag Swapping,” International Conference on iP+Optical Network(iPOP), June 2008.

---

## Chapter 6

---

# 広域イーサネットにおけるTEのための多点間 接続の規模拡張性

### 6.1 概要

これまで第3章から第5章にかけて、二点間を接続する通信技術のTEに注目してきた。本論文の最後の研究である本章では、広域イーサネットにおける多点間接続TEの規模拡張性について扱う。広域イーサネットでは、特にVPN(Virtual Private Network)のようなサービスにおいて複数拠点を結ぶ経路確立技術が求められている。各ネットワークは固有の制御プレーンを保持するオーバーレイモデルを想定し、VLAN IDを指定したパスを確立するためのシグナリング手法を提案した。この手法は実験ネットワーク上に実際に実装し、正確に動作することを確認した。

### 6.2 広域イーサネットにおける多点間接続

第5章にて述べた通り、広域イーサネットはインターネットのための伝送レイヤ技術として注目を浴びている。LANにおける通信技術として開発されたイーサネットは、VLANによりネットワーク上に仮想線サービスを提供する。このVLANの設定をGMPLSで制御することにより、VLANパスの確立自動化が期待されている。ま



た一方で、このVLANのスタック技術であるPBB(Provider Backbone Bridge)および、その拡張であるPBB-TE(PBB Traffic Engineering)が標準化された。広域イーサネットはますます注目を浴びる一方である。

しかし、これらの仮想線サービスはすべてP2P(Point-to-Point)パスに関するものにとどまっている。一方で、Metro Ethernet Forum (MEF)[8]によると、今後のネットワークインフラとして多点間を接続するP2MP(Point-to-MultiPoint)パスの提供が必要とされている。IP-TVや4Kデジタル映画配信などのマルチキャスト型サービスを行う際に、ネットワーク上でトラヒックの複製は、ネットワーク負荷を低減し、効率的な通信を実現することは間違いない。このようなサービスのために、GMPLS制御を拡張し、広域イーサネット上にP2MPパス確立を行うことは非常に重要な試みである。

本研究では、[10]におけるGMPLS制御の広域イーサネットをモデルとして、P2MPのVLANパス確立を行った。RFC 4875[11]は、RSVP-TEメッセージとP2MPシグナリングの手順について定義しているが、シグナリングメカニズムの詳細については定義されていない。そのため、広域イーサネットを制御するP2MPのためのシグナリング手順の詳細を作成する必要がある。

ピアモデルおよびオーバーレイモデルはGMPLSの主要なモデルとして知られている。ピアモデルでは、他の異なるネットワークにまたがる全てのノードは、GMPLSの制御プレーンによって制御される。このモデルの利点は、全てのネットワーク機器が制御ノードによって制御されているため、TEにおける経路の提供が柔軟であるという点である。一方でオーバーレイモデルでは、各ネットワークは固有のコントロールプレーンを保持する。本方式の利点は次の三つである。まず、各ネットワークは自身のネットワークと管理上のモデルを考慮するだけでよい。二つ目に、オーバーレイモデルはネットワーク事業者間で機密性を保持できる。そして三

つ目は拡張性に優れている点である。これは各コントロールプレーンが自身のネットワークドメイン中に存在するノードのみを制御することに起因する。上記の利点から、オーバーレイモデルを本実験に適用することを有効であると考え、本オーバーレイモデルは、ソフトパーマネント接続 (Soft Permanent Connection) を想定している。

RFC 4875 に基づいた P2MP VLAN シグナリングでは、終点ノードにおける出力ポートの指定が必要となる。そこで、オーバーレイモデルのための拡張を施した P2MP VLAN パスシグナリングを提案し実装を行った。本実験は、RSVP-TE シグナリングソフトウェア上にこの拡張を実装し、P2MP VLAN パスがオーバーレイモデルに基づいた実験ネットワーク上で正確に動作することを確認した。

## 6.3 P2MP VLAN パスをサポートした

### GMPLS 制御による広域イーサネットの アーキテクチャ

図 6.1 は、GMPLS 制御による広域イーサネットのアーキテクチャを示している。また、このネットワークはレイヤー 2 VPN サービスを提供する。VLAN パスは GMPLS 制御プレーンにおいて制御ノードの交換するメッセージによって顧客ネットワーク間で確立される。VLAN パス確立のために、二つの手順が必要となる。(1) トポロジ情報を収集し、始点ノードから終点ノードまでの経路を計算 (ルーチング) (2) 経路確立のために全てのスイッチに対し、割り当てられた帯域と VLAN ID を含む経路状態を通知する (シグナリング) 本実験はシグナリング機能に焦点を当てる。

経路は 1 つの始点ノードに対して複数の終点ノード間で確立される。P2MP 経

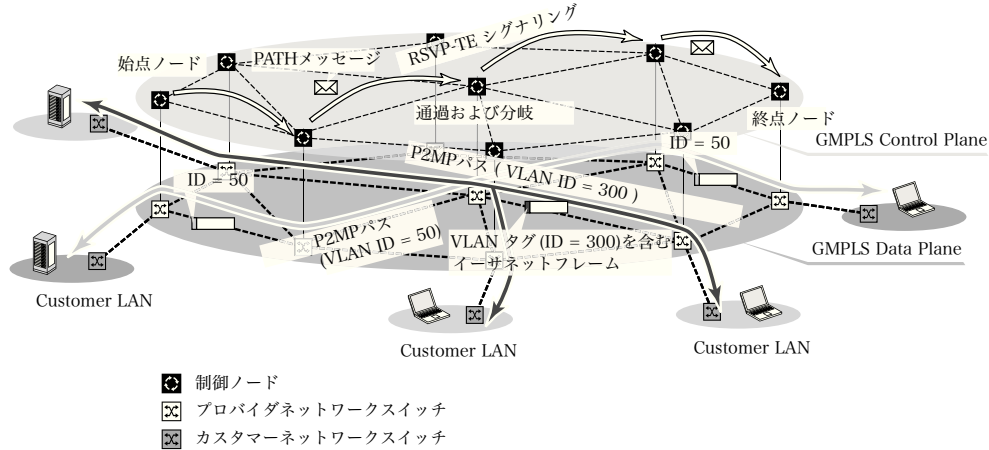


図 6.1: GMPLS 制御における 広域イーサネット

路は P2P 経路のサブセットを保持する。分岐ノードから異なる終点ノードに対して P2P 経路を確立することで P2MP を実現する。始点ノードはシグナリングメッセージを送ることで P2MP 経路を確立する。本メッセージは、各分岐ノードで複製・分散して送信される事で、異なる終点ノードに対して複数 x のリンクを張ることが可能となる。

## 6.4 P2MP VLAN パスシグナリングの実装

### 6.4.1 ピアモデルネットワークにおける P2MP

#### VLAN パスシグナリング

P2MP 経路をサポートするために、シグナリングは分岐ノードから一つの終点ノードに対する sub-LSP を示す Source to Leaf (S2L) sub-Label Switched Path (LSP) を導入する。複数の宛先ノード ID を示すために、P2MP LSP TUNNEL IPv4 (IPv6)

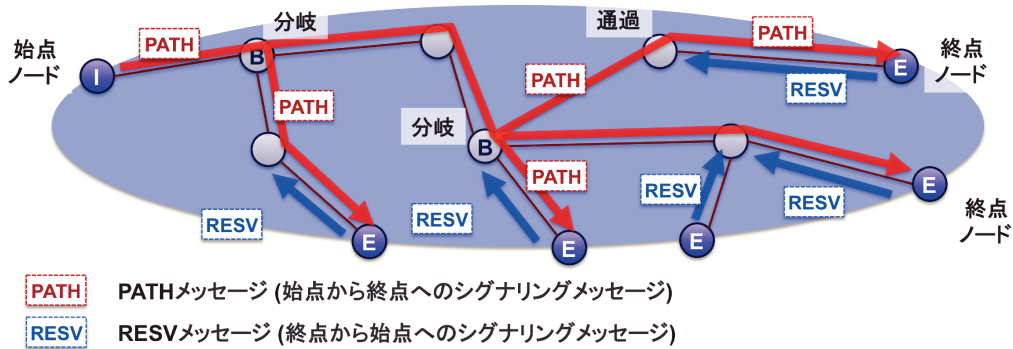


図 6.2: P2MP シグナリング

SESSION オブジェクトと S2L SUB LSP オブジェクトを使用する。しかし、従来型である P2P 用の LSP TUNNEL IPv4 オブジェクトは1つの終点ノード ID を示すための一つのアドレスフィールドしか保持しておらず、P2M LSP TUNNEL IPv4(IPv6) SESSION オブジェクトは、上記のフィールドの代わりに P2MP ID フィールドを保持し、S2L SUB LSP は複数の終点ノード ID を保持する。また、複数の経路を示すために、P2MP SECONDARY EXPLICIT ROUTE オブジェクト (SERO) を使用する。一つの PATH メッセージは、複数の S2L SUB LSP オブジェクトと SERO のペアを含む。

実装のために、分岐ノードにおける処理の詳細を議論させる必要がある。図 6.3 は P2MP VLAN パスシグナリングの例を示している。本例では、始点ノード A が終点ノード D,E,G,H,I に対して P2MP PATH メッセージを送信する。対応する SERO の最初のステップに示されるノードは受け取った PATH メッセージを分岐ノード同様に複製する。PATH メッセージが複製されると、S2L SUB LSP と関係する SERO のペアは各 sub-LSP に転送される。最初の S2L SUB LSP オブジェクトのみが、ペアとしての SERO を持たず、ERO と関連を持つ。他の S2L-SUB LSP オブジェクト

は直後の SERO とペアになっている。図 2 において、ノード B は、二つの SERO が最初のホップとしてノード B を示しているため、2 つのブランチメッセージを作成する。各複製されたブランチメッセージは関連する S2L SUB LSP オブジェクトを継承し、受け取った PATH メッセージから SERO を除く。ノード B は二つのブランチメッセージをノード E と F に転送する。最終的に、ノード B は元の PATH メッセージをノード C に転送する。ブランチメッセージは Sub-Group Originator ID フィールドと P2MP LSP TUNNEL IPv4 (IPv6) SENDER TEMPLATE オブジェクトの Sub-Group ID フィールドによって認識される。上記 2 つの ID はノード ID と、ブランチメッセージを作成する分岐ノードのユニーク ID で設定される。ノード G のように、終点ノードは、他の終点ノードとの間にリンクを保持する場合、ブランチ処理を実行する。

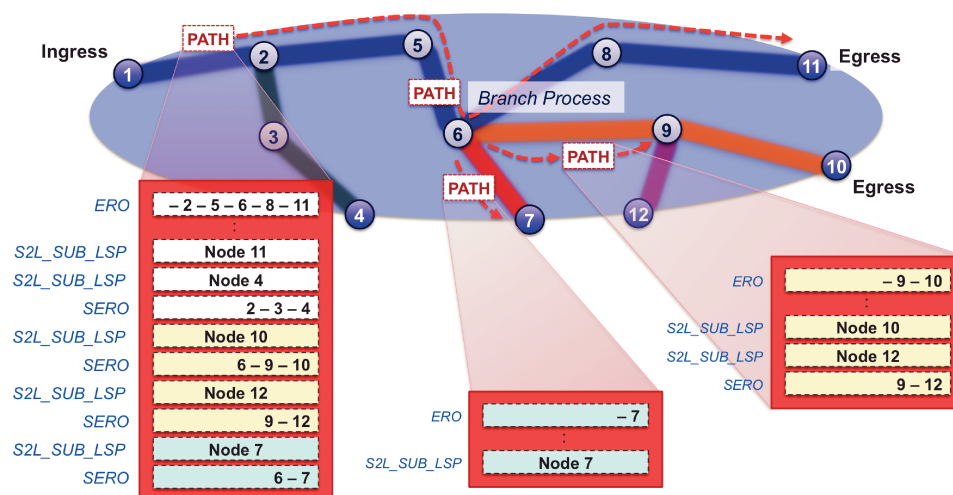


図 6.3: ピアモデルに基づいた標準的な P2MP VLAN パスシグナリングの例

さらに、RFC4875 において VLAN パスについての手法が確立されていないため、P2MP 経路をサポートするために 2 つの拡張が必要である。最初の拡張として、

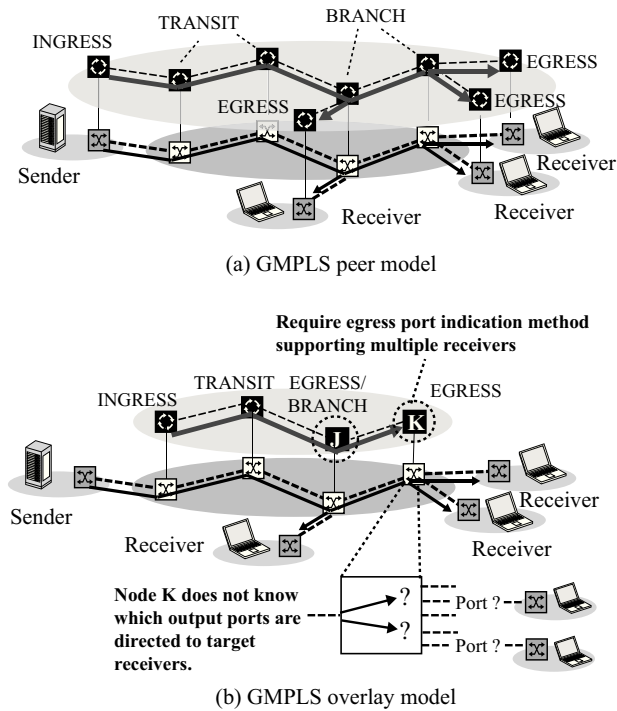


図 6.4: 各モデルにおける P2MP VLAN パスシグナリングの課題

シグナリングプロトコルと組み合わせて VLAN を構成するために、【タグ付き/無し】【ポート番号】【VLAN ID】のラベルフォーマット [12] を採用する。二番目の拡張として、イーサネット経路が双方向であるため、UPSTREAM LABEL を使用することで双方向 P2MP 経路の確立をサポートする。

#### 6.4.2 オーバーレイモデルベース P2MP

##### VLAN 経路シグナリングのための初期拡張

図 6.4 はピアモデルおよびオーバーレイモデルでの P2MP VLAN パスシグナリングを示したものである。ネットワークに図 6.4 (a) のようにピアモデルを配備した場合、プロバイダネットワークのエッジノードではなく、他のネットワークの終点

ノードで終端する。そのため、終点ノードはエッジノードが所属するネットワークに準ずる。

一方で、ネットワークが図 6.4 (b) のようなオーバーレイモデルに配備された場合、プロバイダネットワークのエッジノードで終端する。それ故に、ノード J やノード K のような終点ノードは、標準的な P2MP VLAN パスシグナリングが、終点ノードの上流側ポートのみの VLAN 構成をサポートしているため、どのような下流側ポートで構成するかを認識できない。そのため、出力ポート指定がオーバーレイモデルにおいて必須事項となる。上記のような状況では、エンドツーエンドの接続は [13] のような Soft Permanent Connection であると仮定される。送信元ノードと宛先ノード両方のセットアップは、Permanent Connection として永続的に確立される。そのため、シグナリング手順における送信元または宛先ノードの構成は必要なく、出力ポート表示は直接的に上記の接続性を示す。さらに、P2MP における出力ポート表示は、ノード K のように、ネットワークスイッチを介して同一の終点ノードに接続する複数の宛先をサポートするために必要となる。

実験では、出力ポート指定の機能は複数の宛先をサポートするために導入する。図 6.5 は、オーバーレイのための拡張 P2MP VLAN パスシグナリングの例を示す。始点ノード L が終点ノード J と K に P2MP PATH メッセージを送信する。出力ポート指定のため、ERO の最後のホップに Label sub-object として出力ポート情報を追加する出力ラベル制御 [14] を導入する。各終点ノードは Label sub-object の値を元に、宛先にまたがるポートの VLAN セットアップを構成する。複数の宛先をサポートする拡張は、終点ノードが複数の目的とする宛先を保持する場合、Label sub-object が複数のラベル値を保持することを許容する。例として、図 6.5 におけるノード K は、二つの Label sub-object 値である 10020100 と 10050100 があるためポート 2 とポート 5 において VLAN (ID=100) で出力される。

### 6.4.3 オーバーレイモデルにおける送信元から

#### 始点ノードへの情報交換

既に述べたとおり、オーバーレイモデルにおいてネットワーク管理情報を保護するために、全てのセットアップは送信ノードの代わりに始点ノード(プロバイダエッジノード)で初期化される。LSP 確立において、送信元ノードは始点ノードに、要求LSP、帯域情報、MTU(Maximum Transmission Unit) 値や複数の宛先アドレス等のいくつかの情報と共にリクエストを送信する。上記情報と選択するネットワークモデルによっていくつかのリクエストオプションが存在する。一例として、RSVP-TE シグナリングのためのユーザー網インタフェースとして、Optical Internetworking Forum(OIF)により定義されたメッセージを送信する [15] OIF では複数の Transport Network assigned Addresses(TNA) のリクエストをサポートするために E-line サービスの UNI の手法を指定する。しかし、本指定は現在 P2MP 経路確立をサポートしていない。本実験ネットワークでは、送信元ノードからの LSP 要求はコマンドラインインタフェースを介して送信される。本要求は複数の宛先アドレスと帯域情報を含む。終点ノードにおいて、セットアップメッセージは終端し、宛先ノードには何のメッセージも到着しないが、出力ポート指定はメッセージ中の宛先アドレスから宛先ノードを直接的に示し、宛先ノードは永久的に構成される。

## 6.5 特性評価

図 6.6 は、使用した実験的ネットワーク構成を示している。本ネットワークはオーバーレイモデルに基づいており、9つのノードで構成される。制御プレーンとデータプレーンは分離して構成されている。各ノードは、制御プレーンの GMPLS 制御ノードとデータプレーンのプロバイダネットワークスイッチのペアとして構成



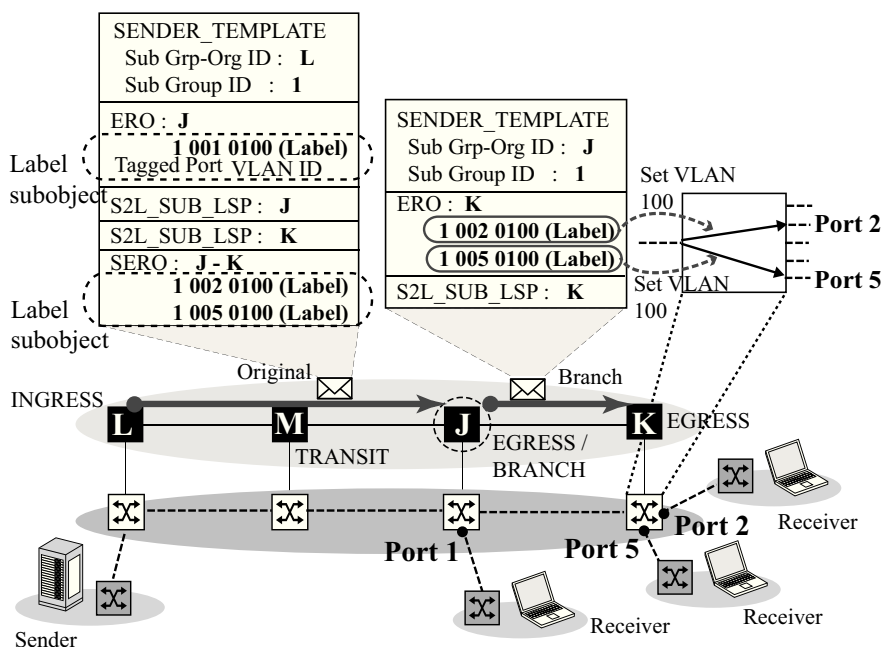


図 6.5: オーバーレイモデルにおける P2MP VLAN シグナリング拡張の例

されている。GMPLS メッセージはアウトオブバンド技術を使用して交換される。簡単のため、エンドユーザーのネットワークスイッチは使用しておらず、送信元ノードと宛先ノードはプロバイダネットワークスイッチと直接接続されている。

Linux ベースの PC と拡張 RSVP-TE プログラムを GMPLS 制御ノードとして使用する。RSVP-TE プログラムは NTT-AT 社製の「GMPLS エンジン」をベースとし、P2MP VLAN パス確立をサポートする拡張を行った。プロバイダネットワークスイッチは、市販の VLAN が使用可能なギガビットイーサネット スイッチ (NEGTEAR 社 GSM7212) を使用した。各イーサネットスイッチは RS-232C シリアルケーブルを介して接続された GMPLS コントロールノードにより制御される。

本実験では、4つの終点ノードと2つの分岐ノードを含む P2MP VLAN パスの確立を行った。始点ノードは 1.1.1.1、終点ノードは 5.5.5.5, 6.6.6.6, 8.8.8.8 そして 9.9.9.9 であり、分岐ノードは 3.3.3.3 と 8.8.8.8 である。LSP セットアップは 1.1.1.1

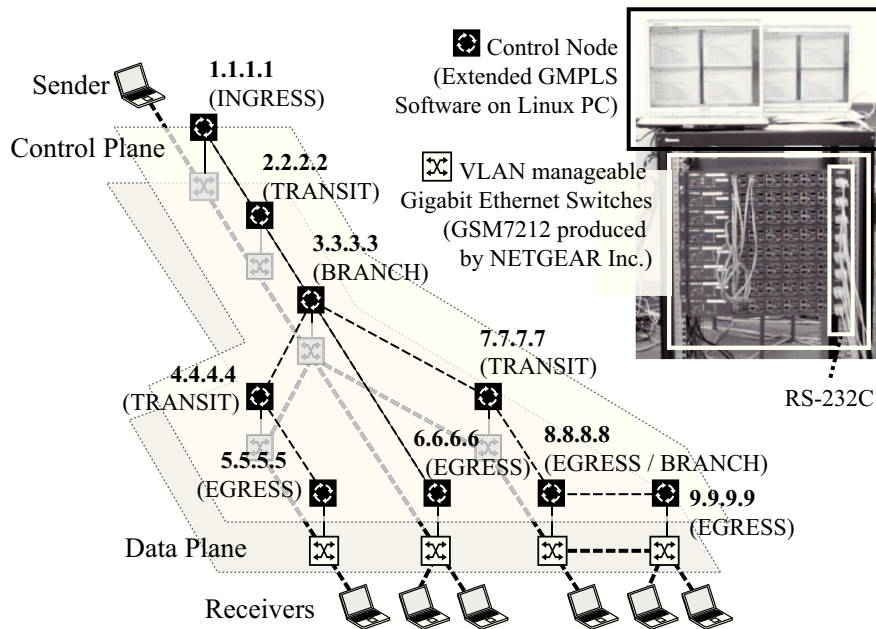


図 6.6: 実験ネットワークの構造

で初期化され、6.4.3で述べたようにコマンドラインインタフェースによって実行される。EROとSEROは、始点ノードによって計算されると仮定した経路を静的に与える。

図 6.7 は、1.1.1.1 が 2.2.2.2 に PATH メッセージを送っている様子を示している。PATH メッセージが含むオブジェクトは、P2MP LSP TUNNEL IPv4 SENDER TEMPLATE オブジェクトと終点ノード 5.5.5.5 と S2L SUB LSP オブジェクトと S2L SUB LSP オブジェクトと終点ノード 6.6.6.6, 8.8.8.8 と 9.9.9.9 の Label sub-object と共に P2MP SECONDARY ERO オブジェクトのペアである。RSVP-TE ソフトウェアは、複数の宛先をサポートする出力ポート表示機能を実装したオーバーレイモデルの拡張である。RSVP-TE ソフトウェアの拡張による自動 P2MP VLAN パス提供の有効性が確認された。

No. .	Time	Source	Destination	Protocol	Info
84	21:40:28.168924	3.3.3.3	4.4.4.4	RSVP	HELLO Message.
110	21:40:32.436658	1.1.1.1	2.2.2.2	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
111	21:40:32.440083	2.2.2.2	3.3.3.3	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
112	21:40:32.443639	3.3.3.3	4.4.4.4	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
113	21:40:32.447408	4.4.4.4	5.5.5.5	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
114	21:40:32.448514	3.3.3.3	6.6.6.6	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
115	21:40:32.450051	3.3.3.3	7.7.7.7	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
116	21:40:32.450771	5.5.5.5	4.4.4.4	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
117	21:40:32.453500	6.6.6.6	3.3.3.3	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
118	21:40:32.454717	7.7.7.7	8.8.8.8	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
119	21:40:32.455172	4.4.4.4	3.3.3.3	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
120	21:40:32.458852	8.8.8.8	9.9.9.9	RSVP	PATH Message. SESSION: IPv4-P2MP-LSP,
121	21:40:32.459976	8.8.8.8	7.7.7.7	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
122	21:40:32.462183	9.9.9.9	8.8.8.8	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
123	21:40:32.462193	3.3.3.3	2.2.2.2	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
124	21:40:32.463792	3.3.3.3	6.6.6.6	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
125	21:40:32.464495	2.2.2.2	1.1.1.1	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
126	21:40:32.467049	1.1.1.1	2.2.2.2	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
127	21:40:32.470204	2.2.2.2	3.3.3.3	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
128	21:40:32.470204	2.2.2.2	3.3.3.3	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
129	21:40:32.471560	3.3.3.3	4.4.4.4	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
130	21:40:32.471784	7.7.7.7	3.3.3.3	RSVP	RESV Message. SESSION: IPv4-P2MP-LSP,
131	21:40:32.473039	4.4.4.4	5.5.5.5	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
132	21:40:32.475790	3.3.3.3	7.7.7.7	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
133	21:40:32.477125	7.7.7.7	8.8.8.8	RSVP	CONFIRM Message. SESSION: IPv4-P2MP-LSP,
184	21:40:34.183367	4.4.4.4	3.3.3.3	RSVP	ACK Message.
185	21:40:34.183496	4.4.4.4	5.5.5.5	RSVP	ACK Message.
186	21:40:34.183870	4.4.4.4	5.5.5.5	RSVP	HELLO Message.
187	21:40:34.184311	5.5.5.5	4.4.4.4	RSVP	HELLO Message.
204	21:40:34.241378	9.9.9.9	8.8.8.8	RSVP	ACK Message. - 59 -
205	21:40:34.241720	8.8.8.8	9.9.9.9	RSVP	HELLO Message.

確立時間 = 40 ms

図 6.7: P2MP シグナリングメッセージ

## 6.6 結論

本論文の最後の研究である本章では、広域イーサネットにおける TE での、多点間接続の規模拡張性について注目した。広域イーサネットでは、特に VPN のようなサービスにおいて複数拠点を結ぶ経路確立技術が求められている。オーバーレイモデルベース GMPLS 制御広域イーサネット上の P2MP VLAN パス確立の詳細について追及し、複数の宛先をサポートする出力ポート識別機能を導入することで、RSVP-TE をオーバーレイモデルに拡張した。そして、拡張 RSVP-TE ソフトウェアによる自動 P2MP VLAN パス提供を実証した。

## 参考文献

- [1] K Fouli and M Maier, “The road to carrier-grade Ethernet,” *IEEE Communication Magazine*, Volume 47, No. 3, Page S30-S38, March 2009.
- [2] E. Mannie (Editor), “Generalized Multi-Protocol Label Switching (GMPLS) Architecture,” *IETF RFC 3945*, October 2004.
- [3] N. Yamanaka, “Photonic Internet Lab.: Breakthrough for Leading Edge Photonic-GMPLS,” *IEICE Transactions on Communications*, Volume E87-B, No. 3, Page 573, March 2004.
- [4] E. Oki, N. Matsuura, K. Shiimoto, and N. Yamanaka, “Bidirectional Path Setup Scheme Using on Upstream Label Set in Optical GMPLS Networks,” *IEICE Transactions on Communications*, Volume E87-B, No. 6, Page 1569-1576, June 2004.
- [5] D. Fedyk, L. Berger, and L. Andersson, “Generalized Multi-Protocol Label Switching (GMPLS) Ethernet Label Switching Architecture and Framework,” *IETF RFC 5828*, March 2010.

- [6] A. Takacs, H. Green, and B. Tremblay, “GMPLS-controlled Ethernet: An Emerging Packet-Oriented Transport Technology,” *IEEE Communications Magazine*, Volume 39, No. 9, Page 118-124, September 2008.
- [7] D. Fedyk, H. Shah, N. Bitar, and A. Takacs, “Generalized Multiprotocol Label Switching (GMPLS) Control of Ethernet Provider Backbone Traffic Engineering (PBB-TE),” *IETF RFC 6060*, February 2011.
- [8] METRO ETHERNET FORUM (MEF) Technical Specification 10.1, “Ethernet Services Attributes Phase 2,” <http://metroethernetforum.org/>, June 2008.
- [9] K. Kikuta, M. Nishida, D. Ishii, S. Okamoto and N. Yamanaka, “Multi-Domain VLAN Path Signaling Method Having Tag Swapping Function for GMPLS-controlled Wide Area Layer-2 Network,” *IEICE Transactions on Communications*, Volume E92-B, No. 4, Page 1353-1356, April 2009.
- [10] K. Kikuta, D. Ishii, S. Okamoto, and N. Yamanaka, “Point-to-Multipoint VLAN path signaling demonstration on the GMPLS-controlled Ethernet test network,” in *Proceedings of OFC/NFOEC 2010*, No. NThE2, March 2010.
- [11] R. Aggarwal, D. Papadimitriou, and S. Yasukawa (Editors), “Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs),” *IETF RFC 4875*, May 2007.
- [12] D. Ishii, K. Kikuta, S. Okamoto, and N. Yamanaka, “An Experiment of Controlling Gigabit Wide Area Ethernet by GMPLS supporting Layer-2 Switching Capability,” in *Proceedings of OFC/NFOEC 2008*, No. OThB5, February 2008.

- [13] D. Papadimitriou (Editor), “Requirements for Generalized MPLS (GMPLS) Signaling Usage and Extensions for Automatically Switched Optical Network (ASON),” IETF RFC 4139, July 2005.
- [14] L. Berger, “GMPLS Signaling Procedure for Egress Control,” IETF RFC 4003, February 2007.
- [15] S. Shew, J. Sadler, “RSVP Extensions for User Network Interface (UNI) 2.0 Signaling,” OIF-UNI-02.0-RSVP, <http://www.oiforum.com/>, February 2008.
- [16] J. Rosenberg, et. al, “SIP: Session Initiation Protocol,” IETF RFC 4003, June 2002.

---

## Chapter 7

---

### 総括

本論文では、インターネット上で規模拡張性に優れたトラフィックエンジニアリングを実現する為の技術に関して提案を行った。各章における要約は次の通りである。第3章では、IP ベースの TE における最適化計算の規模拡張性に注目した。IP ベースの TE では、OSPF 重み付けの最適化により経路を制御するが、この最適化計算は NP 困難でありネットワークサイズとともに計算量が爆発的に増大する。遺伝的アルゴリズムを用いた手法では、計算時間が短縮されるが、それでも大きなネットワークのための最適化では処理時間が十分ではない。計算を高速化するためには、並列プロセッサによる並列処理が有効であるが、従来の方式では各コアの待機時間が大きく、効率的に並列処理が行われなかった。そこで、本提案ではデータの依存関係を解消する事により、効率的に処理を行うアルゴリズムを提案した。16 コアの CPU を用いて計算を行ったところ、従来手法はおおよそ 8 倍までしか高速化されないのに対し、提案手法では 13 倍速く計算した。これらの並列処理に高速化は、今後の大規模並列プロセッサを用いた場合の高速化を期待させる。第4章では GMPLS ベースのマルチメトリック TE における制御システムの規模拡張性に注目した。GMPLS における TE ではメトリックの多様性のサポートが期待されている。従来の GMPLS による LSP 確立では様々なメトリックをサポートした場合、頻繁なメトリックの更新が多大な更新メッセージを発生させ、ノードはその処理に負われてしまう。そこで、フラッドイングを行うルーチングプロトコルを停止し、代わりにシグナリング

メッセージをフラッディングすることによって、メトリック情報の配布を行わずに、最短経路の導出とパス確立を同時に行う FB-RSVP-TE を提案した。この提案手法では、従来手法とほぼ同等のブロック率でありながら、メトリック更新頻度に依存しないメッセージ数でネットワークを制御可能である。これにより、メトリック更新が頻繁なネットワーク上において交換されるメッセージを大幅に削減した。

第5章では広域イーサネットにおける TE のためのドメイン連続性に関する規模拡張性に注目した。マルチドメイン環境の広域イーサネットにおける TE を実現する上で、VLAN が持つ ID 空間の制限はドメインを超えてコネクションの数を制限していた。ドメイン境界ノードにおいてタグをスワップさせるために、そこでシグナリング機能を拡張し、ドメイン毎の ID 空間の独立化を行った。各ドメインでの ID 空間によるコネクション数は制約を持つが、ネットワーク全体では ID 空間の制約から開放されるネットワークを実現し、プロトタイプネットワークを構築する事によってその実現性を検証した。

第6章では、広域イーサネットにおける TE のための多点間接続の規模拡張性に注目した。コネクションレス型である IP ネットワークにおけるマルチキャストの実現は困難であり、GMPLS によるコネクションオリエンテッドネットワーク型のネットワークはマルチキャストの実現のために大きく期待されていた。GMPLS 制御のキャリアイーサネットでマルチキャストで行う場合、VLAN ID の指定が必要となり、これをシグナリングプロトコルで指定する必要がある。そこで、VLAN ID を指定する為の新たなオブジェクトを定義し、プロトタイプネットワーク上に実装を行った。これにより、プロトタイプネットワーク上でイーサネットによるマルチキャストが実現された。

これらの研究成果は、インターネットにおいて TE を実現する上で貢献する。



# 関連著作一覧

## 論文誌

### 博士論文に関連する原著論文

1. Ko Kikuta, Satoru Okamoto, Eiji Oki, and Naoaki Yamanaka, “A Parallelized OSPF Weight Setting Scheme based on a Genetic Algorithm for Multi-Core CPUs,” *Cyber Journals: Journal of Selected Areas in Telecommunications (JSAT)*, Volume 3, Issue 9, No. 3, Page 1-8. Sept. 2013
2. Kou Kikuta, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “Experiment on Point-to-Multipoint VLAN Path Establishment on the overlay-model-based GMPLS-controlled Wide Area Ethernet,” *IEICE Transactions on Communication*, Volume E95-B, No. 10, Page 3302-3306, October 2012.
3. Ko Kikuta, Daisuke Ishii, Satoru Okamoto, Eiji Oki, Naoaki Yamanaka, “Connection setup signaling scheme with flooding-based path searching for diverse-metric network,” *IEICE Transactions on Communication*, Volume E95-B, No. 8, Page 2600-2609, August 2012.
4. Kou Kikuta, Masahiro Nishida, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “Multi-Domain VLAN Path Signaling Method having Tag Swapping

Function for GMPLS controlled Wide Area Layer-2 Network,” IEICE Transactions on Communication, Volume E92-B, No. 4, Page 1353-1356, October 2009.

### その他の論文

5. Sho Shimizu, Wouter Tavernier, Kou Kikuta, Masahiro Nishida, Daisuke Ishii, Satoru Okamoto, Didier Colle, Mario Pickavet, Piet Demeester, Naoaki Yamanaka, “Interoperability Experiment of VLAN Tag Swapped Ethernet and Transmitting High Definition Video through the Layer-2 LSP between Japan and Belgium,” IEICE Transactions on Communication Volume E93-B, No. 3, Page 736-740, Mar. 2010.

## 国際会議発表

### 査読付きの原著論文

1. Kou Kikuta, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “Point-to-Multipoint VLAN Path Signaling Demonstration on the GMPLS Controlled Ethernet Test Network,” Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (OFC/NFOEC) 2010, No. NThE2, March 2010.
2. Kou Kikuta, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “Establishment of Point-to-Multipoint path in GMPLS controlled Wide Area Ethernet,” MPLS 2009 International Conference, No. Wed-1-2, October 2009.

3. Kou Kikuta, Masahiro Nishida, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “Establishment of VLAN Tag Swapped Path on GMPLS controlling Wide Area Layer-2 Network,” Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (OFC/NFOEC) 2009, No. JWA83, March 2009.
4. Kou Kikuta, Masahiro Nishida, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “Implementation of VLAN Path Signaling with Domain Tag Swapping for GMPLS controlled Wide Area Layer-2 Network,” The Conference on the Optical Internet (COIN) 2008, No. C-16-PM1-1, October 2008.
5. Kou Kikuta, Masahiro Nishida, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “GMPLS VLAN Path Establishment using Inter-domain VLAN Tag Swapping,” IP+Optical Network (iPOP) 2008, No. P-2, June 2008.

### 査読付きのその他論文

1. Naohiko Shibuta, Kenta Nakahara, Kou Kikuta, Daisuke Ishii, Satoru Okamoto, Eiji Oki, Naoaki Yamanaka, “Service Composition System Optimizing Network and Service Resources in E3-DCN,” World Telecommunications Congress (WTC 2012), No. PS-21, March 2012.
2. Kenta Nakahara, Ko Kikuta, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, “A Demonstration of IP based Multi-Service Cooperating Cloud Architectures and Technologies for Future Cloud Networking,” World Telecommunications Congress (WTC 2012), No. PS-11, March 2012.

3. Kenta Nakahara, Kou Kikuta, Daisuke Ishii, Satoru Okamoto, Naoaki Yamanaka, "IP based Service Mash-up Provision by uGrid," IP+Optical Network (iPOP) 2011, No. PS-3, June 2011.
4. Satoru Okamoto, Ko Kikuta, Daisuke Ishii, Eiji Oki, Naoaki Yamanaka, "Proposal of the MiDORi GMPLS Traffic Engineering for Energy Optimal Traffic Controlled Networks," MPLS 2010 International Conference, No. Wed-1-2, October 2010.
5. Haruka Yonezu, Ko Kikuta, Daisuke Ishii, Satoru Okamoto, Eiji Oki, Naoaki Yamanaka, "QoS Aware Energy Optimal Network Topology Design and Dynamic Link Power Management," 36th European Conference and Exhibition on Optical Communication (ECOC2010), No. Tu.3.D.4, September 2010.
6. Daisuke Ishii, Kou Kikuta, Satoru Okamoto, Naoaki Yamanaka, "Experiment of Transport and Control Protocols in Control-plane Integrated Next Generation Wide Area Layer2 Network," European Conference on Optical Communication (ECOC) 2009, No. P5.05, September 2009.
7. Daisuke Ishii, Kou Kikuta, Masahiro Nishida, Yutaka Arakawa, Satoru Okamoto, Naoaki Yamanaka, "Experiments of GMPLS based layer2 path control functions for next generation wide area layer2 networks," IP+Optical Network (iPOP) 2009, No. 4-1, June 2009.
8. Satoru Okamoto, Kou Kikuta, Daisuke Ishii, Shuichi Okamoto, Tomohiro Otani, Hideki Otsuki, "Multi ASON and GMPLS Domains Interworking Trials for Automatically Switched Ethernet VLAN Path Provisioning," Optical Fiber Communication Conference and Exposition and The Na-

- tional Fiber Optic Engineers Conference (OFC/NFOEC) 2009, No. NTuB1, March 2009.
9. Teruo Kasahara, Kazumasa Tokuhashi, Kou Kikuta, Daisuke Ishii, Yutaka Arakawa, Satoru Okamoto, Naoaki Yamanaka, “P2P HD Contents Transport Network Based on PLZT High-Speed Optical Slot Switching (OSS),” International Conference on Photonics in Switching (PS) 2008, No. S-02-1, August 2008.
  10. Daisuke Ishii, Kou Kikuta, Satoru Okamoto, Naoaki Yamanaka, “An Experiment of Controlling Gigabit Wide Area Ethernet by GMPLS supporting Layer-2 Switching Capability,” Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (OFC/NFOEC) 2008, No. OThB5, February 2008.
  11. Hideki Otsuki, Daisuke Ishii, Kou Kikuta, Satoru Okamoto, Shuichi Okamoto, Tomohiro Otani, “L2SC Inter-Carrier Interface Interoperability Trial,” MPLS 2007 International Conference, No. Wed-1-4, October 2007.

## 国内学会発表

### 査読なしの原著論文

1. 菊田洸, 石井大介, 岡本聡, 山中直明, “GMPLSシグナリングによる E-tree 確立に向けた RSVP-TE 拡張の実装,” フォトニックネットワーク研究会 (PN), KDDI 研究所, 信学技報. Volume 109, No. 221, PN2009-25, Page 29-34, 2009.10.

2. 菊田洸, 石井大介, 岡本聡, 山中直明, “GMPLS によるギガビット広域イーサネット網の制御実験” フォトニックネットワーク研究会 (PN), 千葉大学, 信学技報, Volume 107, No. 404, PN2007-32, Page 13-17, 2007.12.