

光パスネットワークにおける階層化ルーティングのための 分散ノードクラスタリング手法

福島 行信[†] 原井 洋明^{††} 荒川 伸一^{†††} 村田 正幸[†]

[†] 大阪大学 大学院情報科学研究科 〒 565-0871 吹田市山田丘 1-5

^{††} 情報通信研究機構 〒 184-8795 小金井市貫井北町 4-2-1

^{†††} 大阪大学 大学院経済学研究科 〒 560-0043 豊中市待兼山町 1-7

E-mail: †{y-fukusm,murata}@ist.osaka-u.ac.jp, ††harai@nict.go.jp, †††arakawa@econ.osaka-u.ac.jp

あらまし 光ネットワークにおいてスケーラブルなルーティングを実現するためにはルーティング機構の階層化が必要不可欠である。階層化によって経路表サイズは抑えられるが、経路長は増加する。光ネットワークでは波長連続性制約を満たす必要があり、経路長の増加により光パス設定要求の棄却率は増加すると考えられる。本稿では、ノード間で設定可能な光パス数の最大化を目指してクラスタを構築する手法を提案する。提案手法は、階層化されたネットワーク上で選択される経路として、ファイバ数の大きな経路を用いるようにクラスタを構築することで設定可能な光パス数を増加させると、棄却性能の劣化を抑えられるという考えに基づいている。なお、本手法は、各ノードが分散してクラスタ構築動作を行う分散クラスタリングである。シミュレーションによって、提案手法で構築したネットワークは、経路表サイズが大幅に減少し、棄却率の上昇が抑えられることを示す。

キーワード WDM, 光パスネットワーク, BGP, 階層化ルーティング, 分散クラスタリング

Distributed Clustering Method for Large-Scaled Wavelength Routed Networks

Yukinobu FUKUSHIMA[†], Hiroaki HARAI^{††}, Shin'ichi ARAKAWA^{†††}, and Masayuki MURATA[†]

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

^{††} National Institute of Information and Communications Technology

^{†††} Graduate School of Economics, Osaka University

E-mail: †{y-fukusm,murata}@ist.osaka-u.ac.jp, ††harai@nict.go.jp, †††arakawa@econ.osaka-u.ac.jp

Abstract The scalability of routing table length is indispensable for constructing large-scaled WDM network. Hierarchical routing yields enormous reduction in routing table length, but it also yields increase in path length. In wavelength-routed networks, the increase in path length leads to increase in blocking probability because of less probability to find idle wavelengths on the longer paths. However, if the longer paths have more wavelength resources, we expect that the blocking probability does not get worse. In this paper, we propose a new node-clustering method that maximizes the number of lightpaths between nodes. The key idea of our method is to construct the node-clusters that have much wavelength resources from the ingress border nodes to the egress border nodes, which increases the wavelength resources on routes of lightpaths between nodes. We evaluate blocking probability of lightpath requests and table length by simulation experiments. The results show that our proposed method greatly reduces the table length, while the blocking probability is almost the same as the results with no clustering.

Key words WDM, lightpath network, BGP, hierarchical routing, distributed clustering

1. はじめに

波長分割多重 (WDM: Wavelength Division Multiplexing) 技術を用いた大容量トランスポートネットワークの一つとして、光パスネットワークがある。光パスネットワークでは通信を行う際、まず光パスと呼ばれるチャンネルを設定し、この光パスを用いてデータを送信する。光パスを設定する際には、最初に OSPF (Open Shortest Path First) のようなルーティングプロトコルを用いて経路情報を集め、次に RSVP-TE (Resource Reservation Protocol: Traffic Engineering) や CR-LDP (Constraint based Label Distribution Protocol) のような資源予約プロトコルを用いて経路上で波長を予約する。

これまで、一つのドメイン内 (ポリシーの有効な領域) を対象

とした光パス設定のためのルーティングプロトコルや予約プロトコルが研究されてきた。近年、複数のドメインにわたり光パスを設定するための研究もなされている [1-4]。文献 [1] では、ドメイン間の経路制御機構が満たすべき要件として、独立性を挙げている。すなわち、ドメイン間経路制御機構が用いるルーティングプロトコルとドメイン内経路制御機構が用いるルーティングプロトコルは独立して動作しなければならない。この条件を満たし、かつ、インターネットで広く用いられているドメイン間ルーティングプロトコルは BGP (Border Gateway Protocol) [5] のみである。BGP を例えば文献 [3] で提案されている OBG (Optical BGP) のように光パスネットワークにおけるルーティング/資源予約プロトコルとして拡張することで、ドメイン間光パスの設定に用いることができる。

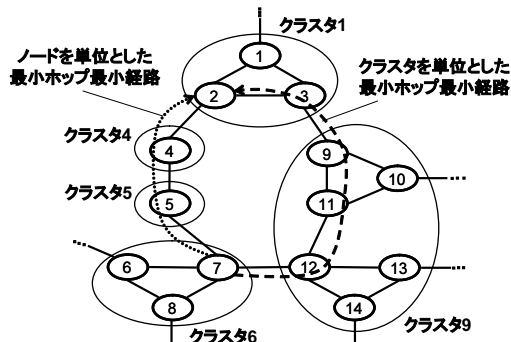


図1 ノードを単位とした最小ホップ経路とクラスタを単位とした最小ホップ経路

一方, BGP の問題点として, 経路数の増大に伴う経路表サイズのスケーラビリティの問題が指摘されている [6]. 経路表サイズの肥大化に対する有効な方法として, 階層化ルーティングがある [7]. 階層化ルーティングを導入したネットワークでは, 複数のノードを 1 つのクラスタとし, クラスタ内経路制御とクラスタ間経路制御を分離する. この場合, 経路表はクラスタ内経路表とクラスタ間経路表の二つに分けられる. クラスタ内経路表はノードを単位とした経路情報を持ち, クラスタ間経路表はクラスタを単位とした経路情報を持つ. このようにクラスタの外部にあるノードの集合をクラスタとして抽象化し, それらのノードへの経路をクラスタへの経路へと集約することにより経路数は減少し, 結果として経路表サイズも減少する.

しかし, 階層化により経路表サイズのスケーラビリティが向上する一方で, 経路情報の集約による経路の変化は光パスの棄却性能の低下につながる. 異なるクラスタに属するノード間に設定される光パスの経路は, それぞれのクラスタの境界ノード (他のクラスタと隣接するノード) を経由する必要がある. その結果, 経由するノード数である経路長が大きくなる. 言い換えれば, クラスタを単位とする最小ホップ経路は必ずしもノードを単位とする最小ホップ経路と一致しない (図 1). ルーティング機構の階層化では, ノード間での光パス設定成功率を高めるようにクラスタを構築することが重要である.

そこで本稿では, 光パス設定要求の棄却率の最小化を目的としたクラスタリング手法を提案する. 本手法では, ノード間で設定可能な光パス数の最大化を目指してクラスタを構築する. 階層化されたネットワーク上で選択される経路として, ファイバ数や波長数の大きな経路を用いるようにクラスタを構築し, 設定可能な光パス数を増加させると, 棄却性能の劣化を抑えられるという考えに基づく.

本提案のクラスタリング手法は, ノードが分散してクラスタを構築する分散クラスタリングである. 従来のクラスタリング手法は, ネットワーク内の特定のノードが最適クラスタを計算し, それを基に実際にクラスタを構築する集中型を想定している. しかし, 大規模光パスネットワークに集中型のクラスタリング手法を適用するのは困難である. なぜなら, 特定のノードがネットワーク全体の大量の情報を保持し, 膨大な時間を費やして計算せねばならないためである. さらに, どのノードがクラスタの計算を担当するべきかという問題も発生する.

以降, 2 章では階層化ルーティング, ノードクラスタリングおよび従来のクラスタリング手法について述べ, 3 章で大規模光パスネットワークを対象とした分散クラスタリング手法を提案する. 4 章ではシミュレーションにより提案手法の有効性を評価し, 最後に 5 章で本稿のまとめと今後の課題を述べる.

2. 階層化ルーティング機構とクラスタリング

2.1 ネットワークモデル

ネットワークモデルを図 2 に示す. ノードはドメインもしくは AS (Autonomous System) に, リンクは光ファイバの集合

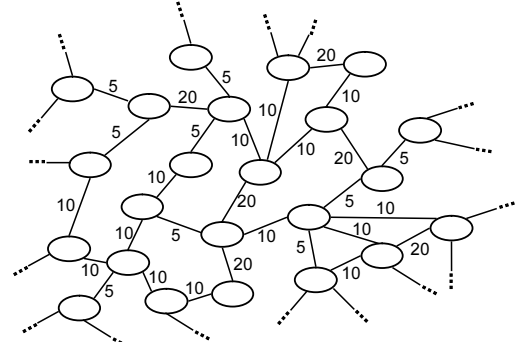


図2 ネットワークモデル

にそれぞれ対応する. 各ドメインの管理主体はそれぞれ異なる. 図 2 の各リンクに付けられた数字は, リンク上のファイバ数を表す.

ドメイン間に設定される光パスは, 各ドメインのゲートウェイに存在するドメイン間光パス制御機構が管理する. 光パスの設定要求がゲートウェイに到着すると, ドメイン間光パス制御機構は, (1) ルーティングプロトコルによる経路情報の収集, (2) 波長予約プロトコルによる空き波長情報の収集, 光パスの設定, 開放を行う. ドメイン間ルーティングプロトコルには BGP のようなパセクトル型のルーティングプロトコルを想定する. これは, BGP がドメイン間光パス設定のためのルーティングプロトコルの要求条件を満たすためである. 以降では, ドメインを表すノードを単にノードと呼ぶ.

2.2 階層化ルーティング機構のためのノードクラスタリング

図 3 に, 階層化ルーティング機構を導入した際に構築されるノードクラスタリングの例を示す. クラスタリングにより作成されたノードの集合をクラスタと呼ぶ. Level x クラスタは, 複数の Level $(x-1)$ クラスタから構成される. Level 0 クラスタはノード自身を表す. 図 3 では, ノード 1, 2, 3, 4, 5 が Level 0 クラスタに対応する.

一つの Level x クラスタに属する Level $(x-1)$ クラスタの数は, 予め決めた上限値までに抑えられる. これはクラスタ内経路表サイズの肥大化を避けるためである. 同様に, Level x クラスタ間経路表のサイズの肥大化も避ける. すなわち, クラスタ数を上限に抑えられない場合は, Level (階層数) を一つ増やしてクラスタ間経路表サイズを抑える. 次章に述べる提案手法は, Level が 2 より大きい階層化へ適用するよう拡張できるが, 本稿では簡単化のため, 図 3 のように階層レベルが 2 までの場合を取り扱う.

2.3 従来のクラスタリングに関する問題

文献 [8] では, コンピュータネットワークを対象としたクラスタリングが研究されており, 最適なクラスタを構築するための問題 (Bounded, Connected, Min-cut 問題) が定式化されている. 以下に Bounded, Connected, Min-cut 問題を示す.

Bounded, Connected, Min-cut 問題

与条件

- 無向グラフ $G = (V, E)$ (各リンク e の重みは $w(e)$)
- クラスタサイズの上限値 $B \in \{1, \dots, |V|\}$

解

- 以下の目的関数を満たすクラスタ集合 V_1, V_2, \dots, V_k .

$$\text{minimize } \sum_e w(e) \quad (1)$$

where $e \in E, e \notin E_i, i \in \{1, 2, \dots, k\} \forall k \in \{2, \dots, |V|\}$

制約条件

- クラスタ V_i 内のネットワークを表すグラフ $G_i = (V_i, E_i)$ は「Connected」
- $1 \leq |V_i| \leq B, \forall i \in \{1, 2, \dots, k\}$

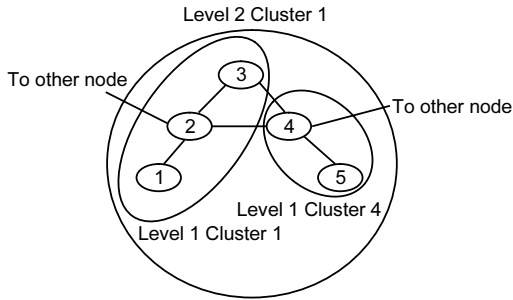


図3 ノードクラスタリングの例

コンピュータネットワークを対象としたクラスタリングの特徴は、各クラスタが「Bounded」かつ「Connected」という条件を要求される点である。「Bounded」は、各クラスタのノード数の上限が B である状態を表し、これがクラスタ内経路表サイズの肥大化を防ぐ。「Connected」は、同じクラスタに属するノードは、同じクラスタに属するノードのみを經由して到達可能な状態を表す。もし「Connected」を満たさなければ、同一クラスタのある2ノードが互いに通信する際に、別のクラスタに属するノードを經由する必要があり、各ノードが他のクラスタに属するノードへの経路も保持せねばならない。この場合、クラスタ内に属するノードへの経路のみ保持することで経路表サイズを抑えるという階層化の目的が達成されなくなる。

Bounded, Connected, Min-cut 問題はNP完全であり、文献[8]では集中型ヒューリスティックアルゴリズムが提案されている。ヒューリスティックアルゴリズムは(1)「Connected」である初期クラスタの作成、(2)クラスタ間でのノード交換、(3)クラスタの合併(Merge)、の3つの部分で構成される。「Connected」である初期クラスタを作成する際には、再帰的二分法(Recursive Bisection)を行う。二分法(Bisection)操作とは、与えられた「Connected」であるノード集合を、二つの「Connected」であるノード集合に分割することである。再帰的二分法操作は、すべてのノード集合の要素数が B 以下になるまで二分法を繰り返す。

再帰的二分法を行うためにはネットワーク全体の情報を保持する必要がある。しかし、大規模なネットワークを対象として再帰的二分法を行う場合には、メモリ量が不足するというスケラビリティに関する問題が発生すると考えられる。また、計算時間も膨大になる。そこで、我々は分散してクラスタを計算するクラスタリングアルゴリズムを提案する。提案手法では、クラスタの境界ノードが、隣接クラスタとクラスタリング操作を行うことにより分散してクラスタを構築する。集中型方式が全体を分割するのに対して、分散型方式では、各ノードを初期クラスタとして合併を繰り返しながらクラスタを構築する。提案手法の詳細は次章に示す。

3. 大規模光パスネットワークを対象としたノードクラスタリング

3.1 光パスネットワークを対象としたクラスタリング問題

1章で述べたように、ルーティング機構の階層化により経路長が大きくなる可能性がある。光パスに使われる波長は経路上ですべて同一である必要があり、経路が長くなるとパスの設定が難しくなる傾向がある。クラスタ構築後のネットワークにおける光パスの経路はクラスタの構成に依存するため、光パス設定要求の棄却率が最小になる経路が選択されるようにクラスタを構築することが重要である。

本章では、以下の3条件を満たすことで、大規模光パスネットワークに適した分散クラスタリングアルゴリズムを提案する。

条件1: クラスタ内/クラスタ間経路表のサイズを一定値以下に抑制する

条件2: 光パス設定要求棄却率を最小化する

条件3: 大規模ネットワークでクラスタを構築可能である

光パスネットワークでは、ノードペア間で設定可能な光パス数の増加が光パス設定要求の棄却性能の低下に繋がるという考えに基づき、クラスタ構築後にノードペア間で設定可能な光パス数の最大化を目指す。我々はこのクラスタリング問題を *Bounded, Connected, Max-path* 問題と呼び、問題を定式化する。

Bounded, Connected, Max-path 問題 与条件

- 光パスネットワーク $G = (V, E)$
- クラスタサイズの上限值 $B \in \{1, \dots, |V|\}$

解

- 以下の目的関数を満たすクラスタ集合 V_1, V_2, \dots, V_k .

$$\text{maximize} \sum_{s=1}^k \sum_{i,j \in V_s} F_{ij} + \sum_{s=1}^k \sum_{i \in V_s, k \notin V_s} F_{ik} \quad (2)$$

ここで、 F_{ij} はノード i からノード j への最短経路上で設定可能な光パス数を表す。 $i = j$ であれば、 $F_{ij} = 0$ となる。

制約条件

- クラスタ V_i 内のネットワークを表すグラフ $G_i = (V_i, E_i)$ は「Connected」
- $1 \leq |V_i| \leq B, \forall i \in \{1, 2, \dots, k\}$

式(2)の目的関数では、全ノードペア間で設定可能な光パス数の最大化を目指す。ノードペア間で設定可能な光パス数は1)同一クラスタ内に属するノードペア間で設定可能な光パス数、2)異なるクラスタに属するノードペア間で設定可能な光パス数、の2種類に区別される。これらのうち、異なるクラスタに属するノードペア間で設定可能な光パス数はクラスタの構成に依存して変動する。なぜなら、經由するクラスタ数を最小とする経路がクラスタ間光パス経路として選択されるためである。

Bounded, Connected, Max-path 問題もNP完全であるため、これに対してヒューリスティックアルゴリズムを提案する。本問題に対する我々の提案手法は、大規模光パスネットワークを対象としたクラスタリングが満たすべき条件(1),(2)を満たす。まず、条件(1)「クラスタ内/クラスタ間経路表のサイズを一定値以下に抑制する」は、2つの制約により、各クラスタが「Connected」かつ「Bounded」という条件にしたがってクラスタを構築するため満たされる。条件(2)「光パス設定要求棄却率を最小化する」は、*Bounded, Connected, Max-path* 問題の目的関数である設定可能光パス数の最大化により実現される。実際に棄却率が抑えられることは4章の棄却率に関する評価により示す。

条件(3)「大規模ネットワークでクラスタを構築可能である」を満たすために、我々はノードが分散してクラスタを構築する分散クラスタリングアルゴリズムを提案する。この場合、各境界ノードはネットワーク全体の情報を保持する必要はないため、三つ目の要求条件が満たされる。3.2節に、提案アルゴリズムで各ノードが保持する情報を示し、続いて3.3節で提案アルゴリズムを示す。

3.2 ノードが保持する情報

図4に、ノードおよび境界ノードが持つ情報を示す。すべてのノードは、(1)ノードクラスタ対応表、(2)クラスタ内経路表、を持つ。これらに加えて、境界ノードは(3)クラスタ間経路表を持つ。以下、それぞれの表の説明と表を参照するタイミングを述べる。

(1) ノードクラスタ対応表

ノードクラスタ対応表は、ノード識別子とクラスタ識別子を保持する。クラスタ識別子は、クラスタに含まれるノードの中のノード識別子の最小値とする。

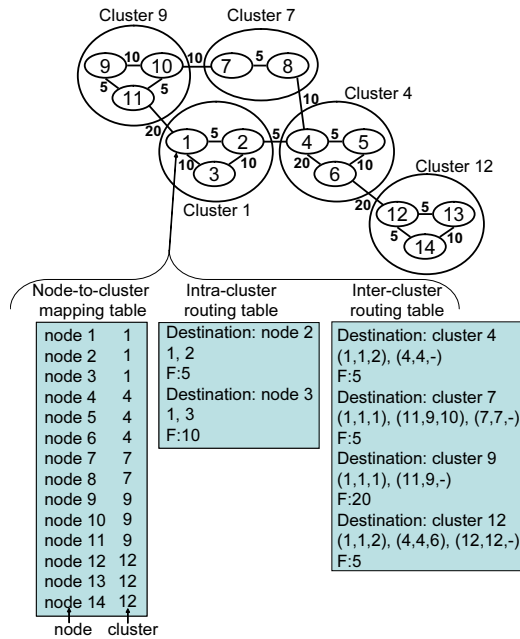


図 4 (1) ノードクラス対照表, (2) クラスタ内経路表, (3) クラスタ間経路表

[クラスタ構築時] 各ノードは (a) 自ノードが属するクラスタの識別子を得るとき, (b) 自ノードが境界ノードであるかを判別するとき, に参照する. クラスタの合併操作後, 各境界ノードは境界ノードでなくなる場合があるが, それは隣接ノードのクラスタ識別子と自ノードのクラスタ識別子と比較することで判断できる.

[光バス設定時] 目的ノードが属するクラスタの識別子を得るためにノードクラス対照表を参照する.

(2) クラスタ内経路表

クラスタ内経路表は, 同一クラスタ内に属する目的ノードまでの最短経路および最短経路上のファイバ数を含む. 図中のノード 2 までの経路情報の中で「1, 2」は経路上のノードリストを表し, 「F: 5」は, ノード 2 までの経路上の各リンクのファイバ数の最小値が 5 本であることを表す.

[クラスタ構築時] 各境界ノードはクラスタ内の他の境界ノードまで設定可能な光バス数を求めるために参照する.

[光バス設定時] ノードが同一クラスタに属するノードへ光バスを設定する際に, 最短経路を得るために参照する.

(3) クラスタ間経路表

クラスタ間経路表は, (a) 目的クラスタまでの最短経路上に存在するクラスタ, および, 各クラスタの入口/出口となる境界ノード, のリスト, (b) 目的クラスタまでの最短経路上のファイバ数を含む. 図中のクラスタ 7 までの経路情報の中で「(1, 1, 1), (11, 9, 10), (7, 7, -)」は経路上のクラスタリストを表す. 各クラスタは (入口境界ノード識別子, クラスタ識別子, 出口境界ノード識別子) と表す. 「F: 5」は, クラスタ 7 までの経路上のクラスタ間リンクおよび中継クラスタ内リンクのファイバ数の最小値が 5 本であることを表す.

[光バス設定時] ノードが他のクラスタに属するノードへ光バスを設定する際に, ノードが属するクラスタのゲートウェイが目的クラスタまでの最短経路を得るために参照する.

なお, クラスタ間経路表は最短経路上に存在するクラスタの入口/出口境界ノードも保持する. これは同一クラスタを経由するクラスタ間経路であっても, 入口/出口境界ノードが異なれば経路上のファイバ数が異なり, それらの経路は互いに別の経路として区別されるためである.

3.3 Bounded, Connected, Max-Path 問題に対する分散クラスタリングアルゴリズム

提案アルゴリズムでは, 各クラスタが隣接クラスタとマージ (合併) 操作を繰り返すことにより適切なクラスタを構築する. マージ操作は二つの隣接するクラスタを合併して一つのクラスタを作成する操作である.

隣接クラスタは 1 つとは限らない. マージ操作の相手を選択する際には, *Bounded, Connected, Max-path* 問題の目的関数である式 (2) の最大化を目指す. 式 (2) の第一項は, クラスタ内では常に物理ホップ最小経路が選択されるためクラスタの構成に依存せず一定の値となる. 一方, 第二項はクラスタの構成に依存してクラスタ間ホップ数最小経路は変動する. よって, 第二項の F_{ik} が増加するようにクラスタを構築することが重要である. このため, 提案手法では各クラスタの入口境界ノードから出口境界ノードまで設定可能な光バス数に着目し, これを最大にするようにクラスタを構築する. その結果, クラスタ間で設定される光バスがファイバ数や波長数の大きなクラスタを経由することになり, F_{ik} が大きくなることを期待している.

提案するアルゴリズム内で用いる記号を以下に示す.

- B : クラスタが含むノード数の上限値.
- R_s : クラスタ V_s 内の境界ノード間で設定可能な光バス数の最小値.
- R_{st} : 隣接するクラスタ V_s と V_t 間のリンク上で設定可能な光バス数の最小値.
- T_w : 隣接クラスタからマージ操作要求が到着するのを待つ時間. 各クラスタは T_w 時間内に送られてくるマージ操作要求に基づいてマージ操作を行う.
- V_{sUt} : クラスタ V_s とクラスタ V_t がマージして構築されるクラスタ.

次に, 提案するクラスタリングアルゴリズムを示す. 以下のアルゴリズムは, 各クラスタ V_i がそれぞれ実行する. ネットワークが階層化されていない場合には, 各ノードがそれぞれ独立したクラスタである状態を初期状態とする. また, ネットワークに新たにノードが接続した際には, そのノードを独立したクラスタとみなす.

- Step 1: T_w を設定し, 隣接クラスタからのマージ操作要求を待つ. T_w が経過すれば Step 2 へ.
- Step 2: 一つ以上の隣接クラスタからマージ操作要求が到着していれば Step 3 へ. 一つもマージ操作要求が到着していなければ, Step 5 へ.
- Step 3: V_i に対してマージ操作要求を送信してきた隣接クラスタ集合 C の各クラスタ $V_t \in C$ 中から, マージ操作後のクラスタ V_{iUt} の境界ノード間で設定可能な光バス数の最小値を求める. 最小値は, $\min(R_i, R_{it}, R_t)$ で与えられる. 次に, この最小値として最も大きな値を送信してきた隣接クラスタ $V_{t'}$ を選択する. $V_{t'}$ 以外のマージ操作要求送信クラスタにはマージ操作を拒否するメッセージを送信し, $V_{t'}$ へマージ操作を受理するメッセージを送信して Step 4 へ.
- Step 4: $V_{iUt'}$ に属するすべてのノードでノードクラス対照情報を更新し, 新たなノードクラス対照情報および新たなクラスタ間経路情報をネットワーク全体へ広告する. $V_{iUt'}$ 内でのクラスタ内経路情報, 境界ノード情報および $R_{iUt'}$ を更新する. 次のマージ操作要求受付終了時間 T_w を設定し, Step 1 へ戻る.
- Step 5: 隣接クラスタの 1 つにマージ操作要求を送信する. 送信先クラスタは次の手順で選択する. ノードクラ

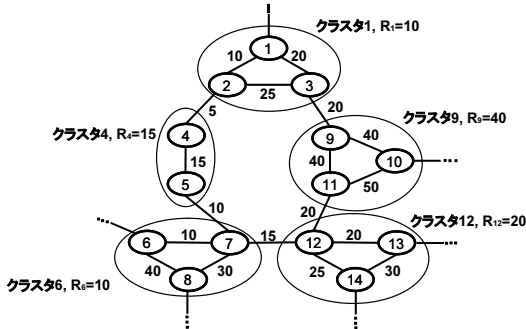


図5 マージ操作前

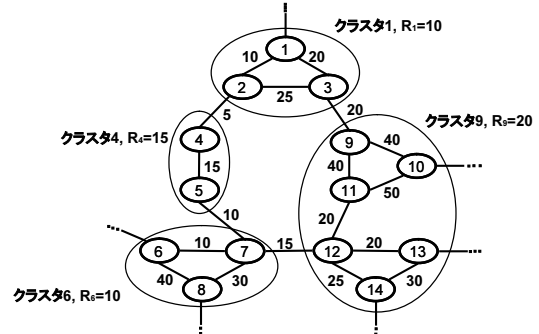


図6 マージ操作後

スタ対応表を参照しクラスタサイズが B 以下になるものが存在すれば、それらの各クラスタ V_t に R_t を問い合わせて受け取る。それらの中から、マージ操作後のクラスタ $V_{i \cup t}$ の境界ノード間で設定可能な光パス数の最小値 $\min(R_i, R_{it}, R_t)$ が最大となるような隣接クラスタ V_t を選択する。上記の手順により選択されたクラスタ V_t にマージ操作要求を送信して Step 6 へ。もしマージ操作後のクラスタサイズが B より大きくなる場合は、Step 7 へ。

Step 6: V_t へのマージ操作要求を受理されたら、Step 4 へ。拒否された場合、次のマージ操作要求受付時間 T_w を設定し、Step 1 へ戻る。

Step 7: マージ操作の対象となる隣接クラスタが存在しないため、クラスタ構築作業を完了する。

Step 5 でマージ操作の相手を決定する際の指標である「マージ操作後のクラスタ内の境界ノード間で設定可能な光パス数の最小値」は、 $\min(R_i, R_{it}, R_t)$ で概算できる。設定可能な光パス数が最小となる境界ノードペアは、1) 二つのノードとも V_i に属する、2) 二つのノードとも V_t に属する、3) 一つは V_i 、もう一つは V_t に属する、のどれかである。1), 2) の場合はそれぞれ R_i, R_t が最小であるためそれぞれが $R_{i \cup t}$ となる。3) の場合には、 V_i に属する境界ノードと V_t に属する境界ノード間のマージ後における最短経路は V_i での境界ノード間経路と V_t の境界ノード間経路を $V_i V_t$ 間リンクで接続して構成される。そのため、その経路上での設定可能な光パス数の最小値 $\min(R_i, R_{it}, R_t)$ が $R_{i \cup t}$ となる。このように各クラスタの入口境界ノードから出口境界ノードまで連続して利用可能な波長数を大きくすることでクラスタホップ数最小経路上で設定可能な光パス数が大きくなり、結果としてエンドノード間で設定可能な光パス数が増大する。なお、 V_i と V_t 間のすべてのリンクがマージ操作後のクラスタの境界ノード間経路として用いられるとは限らないため、マージ操作後のクラスタ内の境界ノード間で設定可能な光パス数の最小値は $\min(R_i, R_{it}, R_t)$ にならないこともある。しかし、マージ操作後のクラスタの境界ノード間で設定可能な光パス数の最小値を正確に求めるためにはマージ操作前の V_i, V_t における各境界ノード間の経路のホップ数をすべて交換する必要がある。よって、ここでは $\min(R_i, R_{it}, R_t)$ を用いる。

図 5, 6 にマージ操作の例を示す。簡単化のため、ここでは波長多重度を 1 とする。図 5 でクラスタ 12 がマージ操作要求を送信する相手を決定する場合、隣接クラスタのそれぞれとマージした際のマージ後のクラスタ内の境界ノード間で連続して設定可能な光パス数を比較する。クラスタ 12 がクラスタ 6 とマージした場合、マージ後のクラスタにおける境界ノード間で設定可能な光パス数の最小値は $R_{12 \cup 6} = \min(R_{12}, R_{12,6}, R_6) = \min(20, 15, 10) = 10$ となる。一方、クラスタ 12 がクラスタ 9 とマージした場合、 $R_{12 \cup 9} = 20$

となる。 $R_{12 \cup 9} > R_{12 \cup 6}$ であるため、クラスタ 12 はクラスタ 9 に対してマージ操作要求を出す。マージ後のクラスタ構成を図 6 に示す。クラスタ 6 からクラスタ 1 までの経路はクラスタ 12 とクラスタ 9 がマージする前は $7 \rightarrow 5 \rightarrow 4 \rightarrow 2$ であるのに対し、マージ後には $7 \rightarrow 12 \rightarrow 11 \rightarrow 9 \rightarrow 3$ へと変更する(經由するクラスタ数が同じ経路が複数ある場合には、経路上で設定可能な光パス数がより大きな経路を選択)。このとき、経路上で設定可能な光パス数が 5 から 15 へと増加する。

4. 評価

4.1 シミュレーションモデル

シミュレーションにより、経路表サイズと光パス設定要求の棄却率を評価する。提案手法(分散 Max-Path)の優位性を示すため、(1) クラスタを構築しない(No Cluster)場合、(2) *Bounded, Connected, Min-cut* 問題を解くアルゴリズムを分散型として拡張したクラスタリング手法(分散 Min-Cut)による結果と比較する。(2) では、クラスタ間リンクのコストにはクラスタ間リンクの数を用いる。各クラスタリング手法では、一つのクラスタが含む要素数の上限値 B を \sqrt{N} (N はネットワーク内のノード数)としている。これは、文献 [7] によりネットワークを M 階層に階層化する際に経路表サイズを最小化する B の値は $\sqrt[M]{N}$ となることが示されており、ここでは $M = 2$ としているためである。マージ操作要求受付時間 T_w は、 $T_w = \gamma \times T$ 、 γ を 1 から 4 までの正の整数の値をとる一様乱数、 T をマージしたクラスタ内で各種情報を交換するために必要な時間より十分大きい 10s とする。

ネットワークトポロジとして、パラメータ α, β をそれぞれ 0.15, 0.2 とした Waxman 手法 [9] から生成したランダムネットワークを用いる。ノード数は 100, 200, 300, 400, 500 とした。ファイバ上での多重化波長数 W は 32 とし、各リンクに配置されるファイバ数 F は 1 から 30 までのランダムな整数とする。各リンクの伝播遅延時間は 10ms とする。

4.2 ノードが保持する経路表の最大サイズ

図 7 に、ノード数が 100, 200, 300, 400, 500 のときの 1 ノードが保持する各種経路表のサイズの最大値を示す。クラスタを構築しない場合は、各ノードは他のノードへの経路を含む経路表のみを保持する。一方、分散 Max-Path(提案手法)と分散 Min-Cut は 3.2 節で説明したようにクラスタ内経路表、クラスタ間経路表、ノードクラスタ対応表の 3 種類の経路表を保持する。経路表サイズの定義は、クラスタ内およびクラスタ間経路表のサイズは各経路のホップ数の総和とし、ノードクラスタ対応表のサイズはエントリ数とする。

図 7 より、分散 Max-Path および分散 Min-Cut はクラスタを構築しない場合と比較して、経路表サイズの最大値を 22% から 33% 程度へと減少させることがわかる。これは、分散 Max-Path および分散 Min-Cut は他のクラスタに属するノードへの最短経路集合を一つのクラスタへの最短経路に集約することで経路数を減らしているためである。ネットワークのノード数が大きくなるほど、経路表サイズを集約できることもわかる。

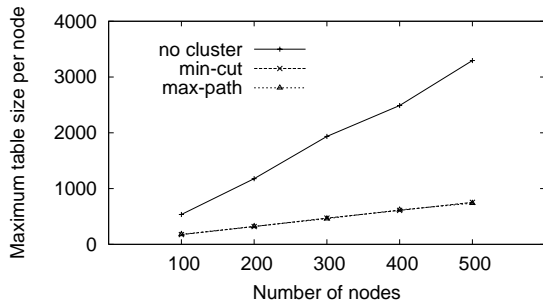


図7 ノードが保持する経路表の最大サイズ

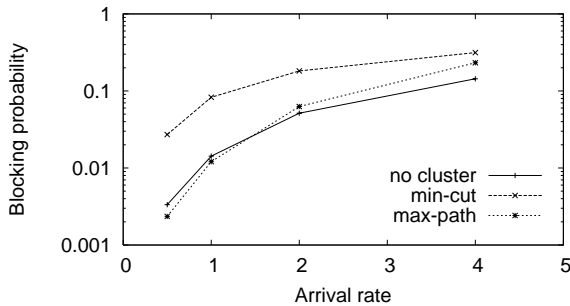


図8 光パス設定要求の棄却率 (平均光パス保持時間:60 秒)

表1 ノードペア間で設定可能な光パス数の平均値

no cluster	min-cut	max-path
307.8	262.6	337

分散 Max-Path 手法と分散 Min-Cut 手法では1 ノードあたりの最大経路表サイズはほぼ等しい。これは、それぞれの手法により構築されるクラスタの数および一つのクラスタが含むノード数がほぼ等しく、その結果クラスタ内/間経路数がほぼ等しくなるためである。

4.3 光パス設定要求の棄却率

クラスタを構築しない場合、分散 Min-Cut 手法によりクラスタを構築した場合、および、分散 Max-Path 手法によりクラスタを構築した場合における光パス設定要求の棄却率を評価する。なお、各クラスタリング手法ではクラスタの構築が完全に終了してから棄却率を評価する。光パス設定要求は、平均到着率 λ (calls/s) のポアソン分布に従って、各送信ノードに様に発生する。光パスの受信ノードはランダムに決定する。光パスの保持時間は平均 60s の指数分布に従うものとする。

図8に光パス設定要求の到着率を変化させた場合の評価結果を示す。ネットワークとしては100ノードのランダムネットワークを用いた。図より、負荷が低いとき(到着率が1以下のとき)には、分散 Max-Path 手法が最も棄却率を抑えられており、クラスタを構築しない場合の結果とほぼ同等の性能を示していることがわかる。これは、分散 Max-Path ではノード間で設定可能な光パス数が大きくなっているためである。表1に各手法におけるエンドノード間で設定可能な光パス数の平均値を示す。表より、分散 Max-Path の効果、すなわち、クラスタを構築する際にクラスタ内で連続して設定可能な光パス数の最小値を最大化することが光パス数の増加につながることをわかる。

負荷が大きくなると、分散 Max-Path とクラスタを構築しない場合で棄却率の大小関係が逆転している。これは、分散 Max-Path 手法によってクラスタを構築した場合における光パスの物理ホップ数がクラスタを構築しない場合よりも大きくなっており、各光パスが使用する波長資源量が大きくなっていくためである。その結果、負荷が高い領域ではより多くのリンクが過負荷状態に達している。各手法における光パスの平均物理ホップ数はクラスタを構築しない場合は3.33、提案手法では4.6、従来手法では4.47となっている。

5. まとめ

本稿では、光バスネットワークを対象とし、階層化ルーティングを実現する分散型ノードクラスタリング手法である分散 Max-Path を提案した。シミュレーションにより分散 Max-Path を従来の Min-Cut の分散拡張、および階層化しない場合と比較した。経路表サイズに関しては、階層化しない場合と比較して最大経路表サイズを22%から33%程度へと減少させることがわかった。ノード規模の増加に対して、経路表サイズの集約がより明確に現われており、ノード数が大きくなればさらに効果が表れることが期待できる。

100ノード規模のネットワークでの光パス設定要求の棄却率に関しては、分散 Max-Path 手法は、分散 Min-Cut 手法と比較して大幅に棄却率を抑えられること、および、負荷が低い場合には階層化しない場合とほぼ同等の棄却率となることがわかった。

スケラビリティに関する問題としては経路表サイズだけでなくノード間で広告される制御情報量が大きくなることも挙げられる。そこで、今後はクラスタ構築時およびクラスタ構築後における制御情報の広告量に関する評価を行う予定である。また、ネットワークに新規ノードが追加された場合には、既存のクラスタ構成を再構築する必要がある。そこで、マージ操作に加えて既存のクラスタ構成を再構築する新たなクラスタリング操作を提案する予定である。

謝 辞

本研究の一部は、文部科学省科学研究費基盤研究(A)(14208027)および文部科学省21世紀COEプログラム(研究開発拠点形成費補助金)の研究助成によるものである。ここに記して謝意を表す。

文 献

- [1] G. Bernstein, B. Rajagopalan, D. Pendarakis, A. Chiu, J. Strand, V. Sharma, D. Cheng, R. Izmailov, L. Ong, and S. Dharanikota, "Optical inter domain routing considerations," *Internet Draft draft-ietf-ipo-optical-inter-domain-02.txt*, Feb. 2003.
- [2] D. Wang, J. Strand, J. Yates, C. Kalmanek, G. Li, and A. Greenberg, "OSPF for routing information exchange across metro/core optical networks," *Optical Networks Magazine*, vol. 3, pp. 34–43, Sept. 2002.
- [3] M. J. Francisco, S. Simpson, L. Pezoulas, C. Huang, J. Lambadaris, and B. St. Arnaud, "Interdomain routing in optical networks," in *Proceedings of Opticomm2001*, pp. 120–129, Aug. 2001.
- [4] C. Pelsser and O. Bonaventure, "Extending RSVP-TE to support inter-AS LSPs," in *Proceedings of 2003 Workshop on High Performance Switching and Routing (HPSR 2003)*, pp. 79–84, June 2003.
- [5] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," *IETF RFC 1771*, Mar. 1995.
- [6] W. Li, "Inter-domain routing: Problems and solutions," *Technical Report TR-128, Department of Computer Science, State University of New York*, Feb. 2003.
- [7] L. Kleinrock and F. Kamoun, "Hierarchical routing for large networks," *Computer Networks*, vol. 1, pp. 155–174, Jan. 1977.
- [8] R. Krishnan, R. Ramanathan, and M. Steenstrup, "Optimization algorithms for large self-structuring networks," in *Proceedings of IEEE INFOCOM'99*, pp. 71–78, Mar. 1999.
- [9] B. M. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications*, vol. 6, pp. 1617–1622, Dec. 1988.