2Q-4

Chord におけるコンテンツ検索の高速化手法の提案

妙中 雄三† 山口 真之介‡ 西野 和典§ 大西 淑雅¶
九州工業大学† 九州工業大学‡ 九州工業大学§ 九州工業大学¶

1 はじめに

近年，P2P ネットワークへの参加ノード数が急激に増加し，スケーラブルな P2P ネットワークに関する研究が注目されている．新たに Distributed Hash Table（DHT）を用いた検索手法が提案され [1]，これらを基にした高速化手法も検討されている [2].

本研究では，Chord[1] のリングを分割した ハッシュ空間を構成し，検索の高速化を目指す．なお本稿では，ノードの増加のみを想定した設計について述べる．

2 設計

本稿の提案手法ではリングの分割処理や，リング間で検索対象のメタデータの同期処理，分割後の Finger Table を維持する処理が必要である．なお，分割後のリング内における検索手法は Chord と同様の手法とした．

2.1 定義

代表ノード

リング分割の判断やメタデータの同期処理を行うノードを代表ノードとする．代表ノードはシステムの制御を担うため，オンライン時間が最も長いノードとする．また，一度決定した代表ノードは離脱するまで変更されることは無いものとした．

リングの識別

分割後のリングを区別するために，本稿では検索を行うリングを検索リングとし，検索リングを識別するためそれぞれにユニークな ID を割り当て，リング ID と呼ぶ．また，Chord のメッセージ全てにリング ID を付加し，同一リングのノードからのメッセージのみを受け取る方式をとる．

2.2 検索リングの分割

（1）新規ノードの参加

代表ノードは分割処理の為に検索リング内のノード数を把握する必要がある．新規ノードは参加する際，初期接続先ノードである bootstrap ノードから代表ノードを取得し，代表ノードに対し参加メッセージを送信する．同時に，リング ID を代表ノードと等しく設定する．その後は Chord と同様の参加処理を行う．

（2）分割処理

検索リングを分割するためには，分割時，分割直後の検索リングにおいて正常な検索が行える必要がある．そのため，検索システム内に存在するメタデータを分割処理と同時に複製することが求められる．また，Chord は全ての検索結果を保証するために，successor pointer が正しい必要があり，各ノードが分割直後にハッシュ空間での次ノードである successor（以降 succ）を正しく認識していなければならない．

図 1 に検索リングの分割処理の例を示す．分割を行う検索リングの代表ノードは，検索リング内のノード数に応じて分割を行う．代表ノードは succ に対して分割時に新たな検索リングへの移動要求を送信し，移動要求を受け取ったノードは succ に対して待機要求を送信する．同様に，succ に対して移動要求，待機要求を交互に送信することで検索リングを一周し，各ノードが分割時にどちらのリングに所属するかを決定できる．

移動･待機要求を送信する際，分割後のリングにおける successor pointer を維持するため，互いに predecessor（以降 pred）, succ を交換する例を図 2 に示す．pred の pred と succ の succ を得ることで，分割後の pred, succ を決定することができる．また，同様にメタデータの送信も行う．各ノードが pred のメタデータを受け取り，管理することで分割後のリング間におけるメタデータの同期を保つことができる．

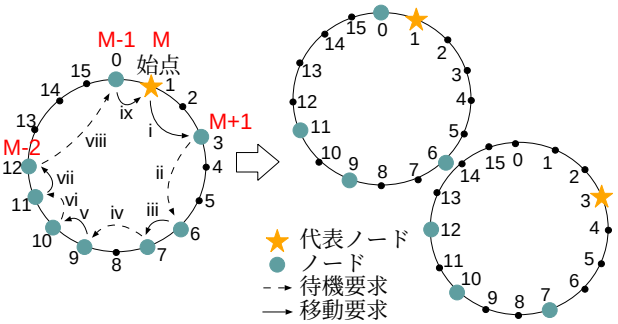

図 1: 分割処理の例

Proposal of a Hi-speed Search Method based on Chord
†Yuzo Taenaka, Kyushu Institute of Technology
‡Shinnosuke Yamaguchi, Kyushu Institute of Technology
§Kazunori Nishino, Kyushu Institute of Technology
¶Yoshimasa Ohnishi, Kyushu Institute of Technology

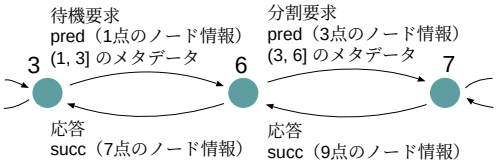

図 2: 分割メッセージの例

なお，検索リング内のノード数が奇数の場合，代表ノードを M とすると，$M-2$，$M-1$，$M+1$ のノードが持つ分割後の情報が正しくない．その際，分割後の succ, pred を $M-2$〜$M+1$ のノード間で調整する．

分割の準備が整うと，代表ノードはリング ID を変更し，通常のリング維持動作を行う．前述の通り，全てのメッセージにリング ID を付加し送信するため，代表ノードのメッセージには新たなリング ID が付加される．他のノードは新たなリング ID のメッセージを受け取ることでリング ID を変更する．

2.3 メタデータの同期

メタデータを同期するために，代表ノード間でメタデータの更新情報を交換する必要がある．そこで，新たに代表ノードで構成する管理リングを作成する．また，各検索リングにおいて検索リングのバージョン管理を行い，更新の度にバージョンを増加させる．

管理リングの各ノードは，同期タスクを定期的に実行し，その時点の状態により（a）検索処理，(b）更新処理のいずれかを実行する．一例として，検索リングが 7 個存在し，管理リングが 4bit ハッシュ空間における各処理の動作を図 3 に示す．

(a) 検索処理

検索処理は，検索リング内でメタデータの更新が無い場合に実行する．管理リングにおいて現在のバージョン番号 +1 を検索キーとして，検索を行う．検索が成功すると該当データ（更新情報）の取得・検索リングへの適用を行い，バージョン番号を増加させる．

(b) 更新処理

更新処理は検索リング内でメタデータの更新が発生している場合に実行する．検索リング内で発生した更新情報を代表ノードに渡し，検索リング内に適用せず削除する．代表ノードは管理リングにおいて現在のバージョン番号 +1 を検索キーとして，検索を行い，更新情報をどのノードが所持するかを検索する．そのノードに更新情報を格納させ，所持する更新情報を削除し適用は検索処理に任せる．

管理リングへの更新情報の格納が衝突した場合は，バージョン番号を再度上げ対処する．また，同期は検索処理に全て任せ，検索リングへの適用順が全ての検索リングで一致させることで，不整合が発生しない．

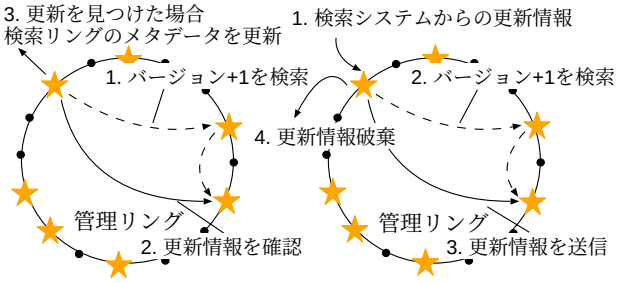

図 3: 左：検索処理，右：更新処理の動作例

2.4 Finger Table の維持

本手法では，検索リングの分割処理後も，succ，pred を除き，分割前の全ての Figner Table を継続して使用する．そこで，他のリングのノード情報を削除するためメッセージに付加されているリング ID が異なる場合，該当ノードを Finger Table より削除する．

また，維持処理においても，メッセージに付加されているリング ID が異なる場合は拒否するため，誤った情報を Finger Table に追加することを防止できる．

3 まとめ

本稿では Chord を基に検索リングを分割することで，P2P ネットワーク上でのコンテンツ検索の高速化手法の提案を行った．本手法では，検索リングを P 分割することで，検索ホップ数を Chord の $O(\log N)$ から $O(\log (N/P))$ へ削減することが見込まれる．検索リングを複数に分割することで，検索ホップ数の減少，検索リング毎のメッセージ数の減少などの利益が得られる．さらに，検索リングを IP レイヤ上の局所性に基づき構成することによる検索速度の向上，検索リングが複数存在することによるメタデータの冗長性の向上が見込まれる．また，検索リング内での検索手法を最適化することでも検索速度の向上が望める．

一方で，分割処理や同期処理において代表ノードに処理を集中させているため代表ノードの負荷軽減や，代表ノードの離脱時における代表ノードの再決定・同期処理の再開，検索リング内のノード数の減少時における検索リングの併合処理を検討する必要がある．

謝辞

本研究の一部は，九州工業大学平成 17 年度研究戦略経費および科学研究費補助金（若手研究（B）16700072）によるものである．

参考文献

[1] Ion Stoica, et al., *Chord: A Scalable Peer-to-peer Lookup service for internet Applications*, ACM SIGCOMM 2001, pp.27-31, August 2001.

[2] Feng Hong, et al., *PChord: improvement on Chord to Achieve Better Routing Efficiency by Exploiting Proximity*, ICDCSW 2005, pp. 806-811, June 2005.