

グループ間通信を抑制する オーバーレイネットワークの構成手法

長尾 洋也[†] 首藤 一幸[†]

1. はじめに

ネットワークの性能向上に伴い、実ネットワーク上に仮想的にネットワークを構築するオーバーレイネットワークと呼ばれる手法が利用されるようになった。オーバーレイネットワークの代表例に Distributed Hash Table (DHT) がある。DHT は、オーバーレイネットワークを利用して連想配列を構成する手法の総称であり、Chord¹⁾ や Kademia²⁾, Tapestry³⁾, Pastry⁴⁾ など、多数のアルゴリズムが提案されてきた。

しかし、これらのアルゴリズムは実ネットワークの情報を考慮せずに DHT ネットワークを構成するため、実ネットワーク上では非効率な経路をとってしまう問題がある。

そこで本研究では、ネットワーク上のノード群を“グループ”に分け、グループ情報を考慮する DHT アルゴリズムの構成手法 *Labeled Flexible Routing Tables (LFRT)* を提案する。

また、LFRT を既存の DHT アルゴリズムである Chord に適用した DHT アルゴリズム LFRT-Chord を設計・実装し、実験を行った。その結果、Chord やラベルを考慮しない場合と比較してグループをまたぐ経路を削減する一方で、経路長の増加が小さくとどまることを確認した。

2. 柔軟な経路表

柔軟な経路表は、DHT アルゴリズムの構成法であり、LFRT のベースとなる手法である。目標分布条件と呼ばれる経路表エントリの ID に関する分布が満たすべき条件を設定し、目標分布条件に近づくように経路表を管理することで経路長を小さく抑える。具体的には、通信などにより入手した経路情報を次々と経路表へ追加し、経路表のサイズの上限を超えたとき、どの経路を削除対象にするか目標分布条件を考慮して選択する。

3. LFRT:Labeled Flexible Routing Table

“LFRT” は、柔軟な経路表をグループ情報を考慮するように拡張した方式であり、ID 情報だけでなく、経路の持つグループ情報を考慮して削除するノードを選択する。“LFRT-Chord” は LFRT に基づいて、Chord を元に構成した DHT アルゴリズムである。

LFRT-Chord では、担当ノードを目標 ID の predecessor としている。これにより、ノードが少ないときワンホップ DHT が実現される。また、目標分布条件として式 (3) を採用する。

$$\delta(x) = \begin{cases} 1 & (\exists e \in X, d(e) = x) \\ 0 & (\text{otherwise}) \end{cases} \quad (1)$$

$$F(x) = \frac{1}{|X|} \sum_{d=1}^{\lfloor x \rfloor} \delta(d) \quad (1 \leq x \leq 2^m - 1) \quad (2)$$

$$F(2x) - F(x) = (\text{一定}) \quad (3)$$

ただし、 X は経路表エントリの ID 集合であり、 m はノード ID のビット長、 $d(x)$ は自ノードから ID x までの距離を示している。

LFRT-Chord は、自身が他のノードと通信を行った際に、通信相手を経路表へ追加する。また、LFRT-Chord では、経路表エントリ数 l が経路表サイズ L を越える場合に経路表エントリの目標分布条件に基づく削除を実行する。

$l = L$ の状態で経路表へノードが追加された場合、 $L + 1$ 個のエントリを一時的に一つの経路表と見なし、 $L + 1$ 個の経路表エントリの中から削除アルゴリズムに基づいて削除対象エントリを決定する。 $X = \{x_0, x_1, \dots, x_{l-1}\}$ として、 $x_{i+1}.pred = x_i$ とし、successor list の長さを s 、 $C(x)$ を式 (4) とする。

$$C(x) = \bar{F}(d(x)) - \bar{F}(d(x.pred)) \quad (4)$$

$$\bar{F}(x) = \frac{1}{m} \log_2 x \quad (5)$$

経路表の最も近い同グループノードより遠くに異グループのノードが存在する場合、そうでないノードを

[†] 東京工業大学 大学院情報理工学専攻 数理・計算科学専攻

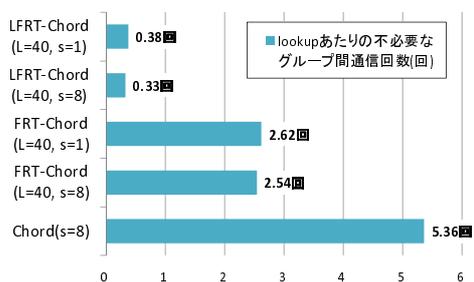


図 1 lookup あたりの unnecessary グループ間通信回数

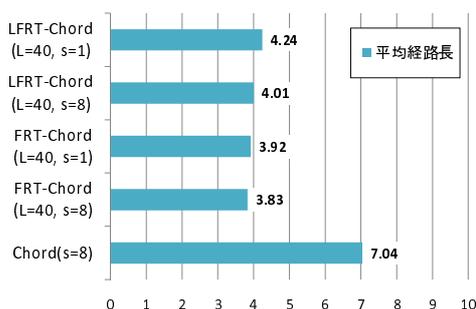


図 2 平均経路長の比較

削除候補から除外する。そして、削除候補のノードの内、 $C(x)$ が最小となるノードを削除対象とする。

4. 評価

グループ情報を考慮しない FRT-Chord および LFRT-Chord をオーバレイ構築ツールキットである Overlay Weaver⁵⁾⁶⁾ を用いて実装し、マシン 1 台上でのエミュレーションを用いて実験した。

- OS : Windows Vista SP2 32bit
- CPU : Intel Core 2 Duo E8400 3.00GHz
- メモリ : 4.00GB
- JRE 6 Update 20
- Overlay Weaver 0.9.9

実験では、10,000 のノードを 1,000 ノードずつ 10 のグループに分け、経路表の構築後の 50,000 回の lookup における経路を集計した。

実験の結果は図 1 および図 2 である。

図 1 は、1 回の lookup において、不要なグループをまたぐ経路を何回利用したかを集計したものである。lookup 元ノードと lookup 先ノードが異なるグループに属する場合、少なくとも 1 回グループをまたぐことになるため、その 1 回は数えていない。このグラフから、グループをまたぐ回数を大きく削減できていることが確認できる。

図 2 は、平均経路長の計測結果である。このグラフ

から、FRT-Chord が不要なグループをまたぐ通信を抑制している一方で経路長を小さく抑えていることが確認できる。

5. まとめと今後の課題

LFRT は、柔軟な経路表の拡張性を利用した構成手法であり、あらかじめ設定したノードグループ情報を考慮した経路表構成を行うことでノードグループをまたぐ通信の抑制を行う。

今後、LFRT を利用した新たな応用手法の検討を行う予定である。

参考文献

- 1) Stoica, I., Morris, R., Karger, D., Kaashoek, M. F. and Balakrishnan, H.: Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications, *SIGCOMM '01: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications*, ACM, pp. 149–160 (2001).
- 2) Maymounkov, P. and Mazières, D.: Kademlia: A Peer-to-Peer Information System Based on the XOR Metric, *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*, pp. 53–65 (2002).
- 3) Zhao, B. Y., Huang, L., Stribling, J., Rhea, S. C., Joseph, A. D. and Kubiatowicz, J. D.: Tapestry: A Resilient Global-scale Overlay for Service Deployment, *IEEE Journal on Selected Areas in Communications*, Vol. 22, pp. 41–53 (2004).
- 4) Rowstron, A. I. T. and Druschel, P.: Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems, *IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*, pp. 329–350 (2001).
- 5) 首藤一幸: Overlay Weaver, <http://overlayweaver.sourceforge.net/>.
- 6) Shudo, K., Tanaka, Y. and Sekiguchi, S.: Overlay Weaver: An overlay construction toolkit, *Computer Communications*, Vol. 31, No. 2, pp. 402–412 (2008).