

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 10 日現在

機関番号：25403

研究種目：基盤研究(C)

研究期間：2010～2012

課題番号：22500064

研究課題名（和文）

頻繁なノード参加における分散ハッシュ表の高速な検索アルゴリズムの開発

研究課題名（英文）

Development of a fast routing algorithm for DHT at frequent joining of nodes

研究代表者

島 和之 (SHIMA KAZUYUKI)

広島市立大学・情報科学研究科・准教授

研究者番号：40263438

研究成果の概要（和文）：

分散ソフトウェアシステムの回復性のため、一定次数のオーバーレイネットワークにおける単純な経路制御アルゴリズムを提案した。このアルゴリズムを用いることによって、P2P型システムのような分散システムにおいて、特定のリソースを持つノードを効率的に検索することができる。提案アルゴリズムの数学的評価によって、平均的な次数がノード数によらず一定であること、および、ノード数に対して経路長が対数となることを示した。次数が一定であることから、ノード数が多い場合でも、ノードの参加や離脱の時、隣接ノードを更新するための負荷が高くなり、スケーラブルなシステムを構築できると考えられる。シミュレーションの結果、Chordアルゴリズムに比べ、提案アルゴリズムにおける損失メッセージ数が少ないことが示された。

研究成果の概要（英文）：

We propose a simple routing algorithm in a constant degree overlay network for resilience of distributed software systems. This algorithm can be used to efficiently locate nodes that provide a particular resource in distributed systems such as peer-to-peer (P2P) systems. Mathematical evaluation of the proposed algorithm shows the expected degree is constant not depending on the number of nodes, and the expected path length is logarithmic with the number of nodes. The proposed algorithm is scalable because the degree is constant, that is, the load of updating neighbors does not become high when the number of nodes becomes large. A simulation shows the number of lost messages for the proposed algorithm is less than that of Chord algorithm.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	700,000	210,000	910,000
2011年度	1,000,000	300,000	1,300,000
2012年度	500,000	150,000	650,000
年度			
年度			
総計	2,200,000	660,000	2,860,000

研究分野：総合領域

科研費の分科・細目：情報・計算機システム・ネットワーク

キーワード：オーバーレイネットワーク

科学研究費助成事業（科学研究費補助金）研究成果報告書

1. 研究開始当初の背景

分散ハッシュ表(Distributed Hash Table, DHT)とは、通信ネットワークに接続した不特定多数のノードが分担して記憶するハッシュ表であり、次の機能を持つものである。

- キーの集合を分割し、ノードに割り当てる。
- ノードの参加や離脱のとき、ノードに割り当てるキーの集合を更新する。
- 指定されたキーの担当ノード（そのキーを割り当てられたノード）を検索する。
- 指定されたキーの担当ノードが、そのキーに対応する値を登録または取得する。

DHT の評価指標として、経路長と次数がある。経路長(path length)とは、2つのノード間の経路の長さである。担当ノードを短時間で検索するためには、経路長が短いほど良い。ただし、一般の有向グラフ上の最短経路を求める方法(e.g. ダイクストラ法)は、全ての有向辺を知る必要があるため、DHTには適さない。次数(degree)とは、各ノードの隣接ノードの数である。ノードの参加や離脱のとき、隣接ノードを短時間で更新するためには、次数が少ないほど良い。隣接ノードが更新されていないと、担当ノードの検索において経路長が長くなるため、検索に時間がかかり、検索に失敗することもある。代表的な DHT として、Chord[1]と CAN[2]が挙げられる。Chord の次数は $O(\log n)$ 、経路長は $O(\log n)$ である。ここで、 n はノード数である。CAN の次数は $O(d)$ 、経路長は $O(dn^{1/d})$ である。ここで、 d は次元数である。

[1] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker: A Scalable Content Addressable Network, Proc. SIGCOMM 2001, pp. 161-172 (Aug. 2001).

[2] I. Stoica, R. Morris, D. Karger, M.F. Kaashoek, and H. Balakrishnan: Chord: A scalable peer-to-peer lookup service for internet applications, In Proc. ACM SIGCOMM 2001, pp. 149-160 (San Diego, Aug. 2001).

2. 研究の目的

本研究では、提案するアルゴリズムを実装し、複数のコンピュータによる実験を行い、他の DHT との比較評価を行う。不特定多数のノードが随時、参加、離脱、検索を行う状況を想定する。評価基準として、次数と経路長だけでなく、ノード間でやり取りされるメッセージの数、担当ノードを検索するまでにかか

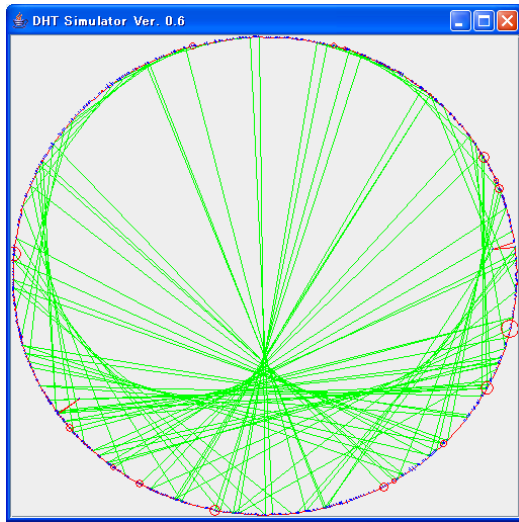
る時間なども計測する。ノードの参加と離脱に比べて検索の頻度が多い場合や少ない場合、ネットワークの遅延や障害がある場合などにおける DHT による特徴の違いなどについても明らかにする。

3. 研究の方法

分散ハッシュ表では、あるキーに対する担当ノードを検索するためにオーバーレイネットワークを用いる。オーバーレイネットワークを構成するために多数のアルゴリズムが提案されており、アルゴリズムがオーバーレイネットワークの性能に影響する。オーバーレイネットワークのアルゴリズムを評価するためには、実際に通信を行う実験や仮想的に通信を行うシミュレーションによる方法などがある。実験による方法では、より現実に近い結果を得られると期待されるが、多数のノードからなるオーバーレイネットワークを構成することは難しい。シミュレーションによる方法では、多数のノードからなるオーバーレイネットワークを構成することは可能であるが、その結果が現実に近いことが重要である。

本研究では、実験とシミュレーションにおいて同一コードでアルゴリズムを実装可能な OverlayWeaver を使い、PlanetLab 上の実験によって得たデータとシミュレーションによって得たデータとを比較した。その結果、統計的に有意な差があることを確認し、第9回情報科学技術フォーラムにて発表した。データの差の原因として、マルチスレッドによる実時間シミュレーション方式であったことが考えられる。マルチスレッドを用いると、1つのCPUで複数ノードの処理をシミュレートできる。しかし、ノード数が多くなるにつれてCPUの負荷が高くなり、1つのCPUで1つのノードを処理する実験と比べて処理時間の差が大きくなる。

そこで、単一スレッドによる離散イベントシミュレーション方式を採用し、実験とシミュレーションにおいて同一コードでアルゴリズムを実装可能なシミュレータを新たに開発する。開発したシミュレータに、提案アルゴリズムと代表的な従来手法である Chord のアルゴリズムを実装する。不特定多数のノードがランダムに参加、離脱、検索を行う状況をシミュレートし、各メッセージの送信時刻、送信ノード、受信ノード、宛先キー、送信源からの経路長をログに記録する。このログを解析することによって、メッセージの経路長や到達率を求める。



4. 研究成果

開発したシミュレータは Java 言語で Chord アルゴリズムが 514 行、提案アルゴリズムが 705 行、その他の共通部分が 2429 行であった。開発したシミュレータでノードとメッセージを可視化したものを上に示す。大きな円周上の小さく赤い円がノードを示す。ノードがメッセージを送信した直後に赤い円は一瞬大きくなり、次第に小さくなる。緑の線はメッセージの経路を示す。

開発したシミュレータと OverlayWeaver を PlanetLab 上の複数のノードで実行した場合と 1 台の PC で多数のノードをシミュレートした場合とを比較した。その結果、開発したシミュレータの方がメッセージの遅延時間の誤差が小さいことが分かった。この結果について[7]で発表した。

開発したシミュレータを用いて、Chord と提案アルゴリズムを比較する実験を行った。イベントは、ノードの参加と離脱、メッセージの送信と受信である。ノードの参加と離脱は、平均 10 秒の指数分布に従うランダムな時間間隔で起こる。ランダムに選ばれたノードがランダムに選ばれたキーへ Lookup メッセージを送信する。これを 1 秒間隔で 1 万回行った。下位のネットワークのプロトコルは UDP のようなコネクションレスと仮定した。メッセージの送受信に要する遅延時間は平均 50ms の指数分布に従う。Chord と提案アルゴリズムは再帰型で実装した。提案アルゴリズムの定数 β は 8 とした。ノード数は 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536 とした。

予想に反して、提案アルゴリズムの方が Chord よりも平均経路長が若干長くなる結果となった。原因として、メッセージの経路に

誤りがあることが考えられた。そこで、FALSE:メッセージがキーの担当ノード以外に届く誤りと LOST:メッセージがどのノードにも届かず損失する誤りの数を計測した。結果を下の表に示す。

FALSE の数

N	Chord	Proposed
512	98	51
1024	23	20
2048	15	19
4096	7	6
8192	9	2
16384	2	0
32768	0	2
65536	1	0

LOST の数

N	Chord	Proposed
512	1572	728
1024	691	371
2048	443	192
4096	235	110
8192	96	60
16384	66	29
32768	27	15
65536	24	10

FALSE の数はどちらのアルゴリズムでも 1%以下であったのに対し、LOST の数は Chord の場合 16%、提案アルゴリズムの場合 7%と開きがあった。一般に経路が長いほど、LOST が生じやすい。LOST が生じると、経路長が求まらず、平均経路長の計算に加えられない。このため、LOST が多い Chord の平均経路長が短くなったと考えられる。公平な比較を行うためには、経路長よりも LOST 数の方が適していると考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 0 件)

[学会発表] (計 7 件)

[1] H. Kwon, K. Shima, Y. Sato, and M. Ohba, A simple routing algorithm in a constant degree overlay network, International Symposium on Software Reliability Engineering, Student papers, Session #4-1, Nov. 2011.

[2] 田丸純、阿部紘一、島和之、ICカードとオーバーレイネットワークによる災害時の安否確認システ

ム, 電子情報通信学会 2012 年総合大会, 2012 年 3 月 20 日.

[3] 手嶋達也, 為岡未来, 島和之, リング型オーバーレイネットワークにおけるルーティングアルゴリズムの提案, 情報処理学会ソフトウェア工学研究会, 2012 年 3 月 16 日.

[4] 田丸純, 阿部紘一, 島和之, 前田香織, オーバレイネットワーク上に構築した安否確認システムの有効性に関する実験的評価, 第 5 回インターネットと運用技術シンポジウム (IOTS2012), 2012 年 12 月.

[5] 橋本淳志, 権孝眞, 伊藤和也, 島和之, 災害時を考慮した広域分散環境におけるサイト選択方法の考察, 情報処理学会全国大会, 2013 年 03 月.

[6] 田丸純, 橋羽里沙, 島和之, 寺田英子, 前田香織, ICカードと衛星通信を用いた安否確認システムの実験的評価, 情報処理学会全国大会, 2013 年 03 月.

[7] 手嶋達也, 島和之, オーバレイネットワークのルーティングアルゴリズムの離散イベントシミュレーション, 情報処理学会全国大会, 2013 年 03 月.

〔図書〕 (計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

なし

6. 研究組織

(1) 研究代表者

島 和之 (SHIMA KAZUYUKI)

広島市立大学・

情報科学研究科・准教授

研究者番号: 4 0 2 6 3 4 3 8

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: