

10の10乗規模の個品追跡を可能とする トレーサビリティシステム向けID解決機構

土井 裕介^{†1} 若山 史郎^{†1} 石山 政浩^{†1}
尾崎 哲^{†1} 井上 淳^{†2}

電子タグを用いた技術の目標の1つに、物品1つ1つに対してIDを割り当て、それらの物品に対応する情報をIDと紐付ける、地球規模で利用できる個品トレーサビリティシステムが存在する。本研究では、著者らが提案してきた「規模拡張可能な個品レベルのトレーサビリティシステム」に向けたID解決機構の設計と実装について、規模拡張性の評価を行う。提案システムはトレーサビリティシステムの「ID→情報」の紐付けの機能を「ID→DBの位置」の紐付けと「DBの位置→情報」の紐付けとの2段に分割した。前段の個品ID解決サービスには分散ハッシュテーブルを用い、規模拡張可能な個品レベルID解決機構とする。そのうえで、商品カテゴリごとの個品ID解決サービスをDNSと結合するDHT-DNSマウンタを導入する。DHT-DNSマウンタにより、多様な性質の多数のカテゴリを持つ商品IDに対して個々に適切な解決機構を提供できる大規模個品トレーサビリティシステム向けID解決機構が実現可能であることを示した。また、現状の実装の延長線上で達成可能な性能を確認した。

An ID Resolution Scheme to Enable 10¹⁰ Order Item-level Traceability System

YUSUKE DOI,^{†1} SHIROU WAKAYAMA,^{†1} MASAHIRO ISHIYAMA,^{†1}
SATOSHI OZAKI^{†1} and ATSUSHI INOUE^{†2}

One of the promised goals of RFID based technologies is a global-scale item-level traceability system to bind each item and its information. In this paper, we evaluate scalability of the ID resolution scheme of our scalable item-level traceability system. The proposed scheme divides “ID→information” retrieval in traceability system into two stages of “ID→DB location” and “DB location→information.” The first stage is item-level ID resolution service per item class and we apply distributed hash table (DHT) to make the service scalable. In addition, we introduce the DHT-DNS mounter that unites each per-class DHT resolution service under a domain under the DNS. The authors show DHT-DNS mounter enables a large-scale item-level traceability system and estimated performance of our implementation.

1. 背景

トレーサビリティシステムは、主に物品の現在の所在、流通経路、および付加価値情報を提供する。一般的には、物品にIDを付与し、これを鍵に有用な情報を得ることを目的としている。トレーサビリティシステムの一例として、EPCglobalによって提唱されているアーキテクチャ¹⁾をあげる。これは、電子タグを中

心にしたトレーサビリティシステムであり、EPC-ISと呼ぶ情報サービスとONSと呼ぶID解決機構、タグ自身の標準やタグ内部に記載するデータ形式の標準²⁾などからなる。

トレーサビリティシステムは、ある管理領域に閉じたトレーサビリティシステムと、オープンなトレーサビリティシステムに分類できる。たとえば、重要ドキュメント管理のためのファイリングシステム³⁾のように、域内に閉じたトレーサビリティシステムにおいては、個品1つ1つを追跡することに量的な困難はない。一方、世界中からのアクセスを許す、オープンなトレーサビリティシステムにおいては、データベースの位置とタグIDの組合せの数を考慮しなければならない。規模拡張性の制約から、商品情報提供の応用例

^{†1} 株式会社東芝研究開発センター通信プラットフォームホームラボラトリー
Communication Platform Laboratory, Corporate R&D
Center, TOSHIBA Corporation

^{†2} 株式会社東芝イノベーション推進部
Innovation Promotion Division, TOSHIBA Corporation

においては特定系列の商品のみを個品追跡するシステムを構築する事例^{4),5)}がほとんどである。

「全世界の個品 ID」のようなフラットで莫大な規模の ID 空間を管理するために、Chord⁶⁾などの分散ハッシュテーブル (DHT) を利用するのは自然な考え方である。実際、特定の業者の個品 ID を管理するためだけであれば、大きな問題はないように考えられる。しかし、システムが普及するにつれ、単一の DHT で多くの業者の多様な商品カテゴリを取り扱うことを想定すると、DHT の一様性が問題となる可能性がある。

ここでいう DHT の一様性は、ハッシュ関数によるノード選択方式に由来する。DNS のような木構造の名前解決システムであれば、異なる実装やパラメータを導入した名前解決サービスを同じ名前空間において実現できる。一方、DHT を用いた ID 空間ではハッシュ関数により個々の名前がハッシュ値の空間全体に分布してしまう。このため、ID 空間の一部に対する取扱いを区別することが原理的に困難で、ID 空間全体が一様な性質を持つこととなる。

本研究で考える DHT の性質には、おおきく分けて 2 つの軸を想定する。1 つはコストと性能のバランスであり、もう 1 つはコストと信頼性のバランスである。本研究ではコストと性能のバランスのうち、特にコストと問合せに必要な時間のバランスについてのみ取り扱う。なお、ここでいうコストとは、一定の性能を実現するのに必要なハードウェア量やトラフィック量である。加えて、近傍ルーティング方式の有無、登録データの永続性の有無などといった相違が考えられる。

また、商品カテゴリには、商品のライフサイクルが短く商品 ID の数が膨大になるカテゴリ (食品など) と、逆にライフサイクルが長く商品 ID の数が限定的なカテゴリ (車などの耐久消費財)、あるいは要求の季節変動の激しい商品 (お正月用品や学習用具) などを含む。ここで、商品カテゴリの性質が ID 空間に求める性質、特にコストへの許容条件が異なる場合、単一の DHT システムではそれらを吸収できない。

そのうえ、単独の DHT で全個品の情報を取り扱う場合、その DHT が動作しなかった場合の責任を誰がとるのか、コスト負担はどのようにするのか。一様な DHT にすべての ID 解決を任せると、責任分界点を明確化できない。

そこで、物品の ID と、それぞれに対応する全世界に分散した情報との柔軟な紐付けを目指して、著者らは小規模から円滑に拡張でき、個品に対する現場レベルの情報を管理可能なトレーサビリティシステム⁷⁾を提案している。特に、その中で用いる DHT と DNS

を併用するハイブリッド ID 解決機構⁸⁾は、タグ ID から関連情報を持つ DB の URL を解決するための ID 解決機構である。これは、DHT を用いた商品カテゴリごとの個品 ID 解決サービスと、異なる個品 ID 空間を DNS の名前空間に結合するための DHT-DNS マウンタと呼ぶ機構からなる。また、本システムを具体的に適用したフィールド実証の事例は、文献 9) にまとめられている。

本研究では、トレーサビリティシステムに対する応用を主眼とし、複数の商品カテゴリに由来する異なる性質の個品 ID 空間を統合でき、かつ、個品 ID を取り扱うのに十分な規模拡張性を持つ ID 解決機構を構築し、評価する。さらに、DHT-DNS マウンタにより ID 解決機構を結合することで、多様な商品カテゴリそれぞれにあわせて適切な性質の個品 ID 空間を提供できることを示す。量的には、DHT を用いた ID 解決機構が、e-Japan 重点計画 2004¹⁰⁾などで提示されている量的目標である 100 億個 (10^{10}) 規模のタグ ID を 0.3 秒で取り扱うのに十分な規模拡張性を備えることを示す。このため、特に問合せ時間を短縮するための工夫により、個品 ID 空間の性質の差別化を行う。なお、この時間は、「電子タグ高度利活用に関する研究開発」に係る委託研究において設定された電子タグ情報解決の目標時間である。

2. 関連研究

EPCglobal のアーキテクチャでは、タグ ID はヘッダ部 (Header) を除くと、生産者 ID (General Manager)、商品カテゴリ ID (Object Class)、個品 ID (Serial Number) の 3 階層からなる ID を持つ。一例として、GID-96 と呼ばれるタグデータ形式を、図 1 に示す。

ONS¹¹⁾は、この 3 階層のうち上位 2 階層を用いた ID 解決を行う、EPCglobal で標準として定義されている ID 解決機構である。DNS の機構をそのまま用い、電子タグの ID から情報サービス (EPC-IS) の URL を解決する。一方、ONS は DNS の限界に束縛される。シリアル番号のような多量で構造を持たないものを効果的に分散させる機能は、DNS には存在しない。つまり、ONS 単体は、DNS の名前空間木を持つ性質から、複数の性質の異なる商品カテゴリ向け ID

Header	General Manager	Object Class	Serial Number
8 bit	28 bit	24 bit	36 bit

図 1 GID-96
Fig. 1 GID-96.

空間を構築できる。しかし、その空間に個品 ID レベルの情報を含むことはできない。

これに対して、本研究では個品 ID 解決に DHT と DNS を組み合わせて用いる。DHT を用いた名前解決機構として、Cox らの研究¹²⁾ が知られている。Cox らの研究では、DNS を DHT で置き換える場合の困難さがいくつか指摘されている。そのうちの 1 つに、異なる性質の個品 ID 空間を同一のシステムで実現できないという問題がある。つまり、DHT を利用した個品 ID 空間単体では、ID 空間を構築するのに十分な規模拡張性を持つ可能性がある一方、様々な商品カテゴリが相反する性質の個品 ID 空間を要求した場合、これを満たすことができない。

本研究の貢献の 1 つは、個々の商品カテゴリごとに異なる特性の個品 ID 空間を用意できるような ID 解決システムを提供することである。ただ、現時点では本システムに組み込んだ DHT アルゴリズムは Chord⁶⁾ しかない。そこで、Chord アルゴリズムにパラメータを導入することで特性の 1 つの軸、経路維持にかかるトラフィック量と問合せ解決にかかる時間 (ホップ数) のトレードオフを制御可能にした。先行事例としては、Alima らによる DKS¹³⁾ があげられる。Alima らの DKS は、冗長性と探索の深さをパラメータとして与えられる DHT の 1 群である。DKS の探索空間は、1 階層あたり k 分割された階層が L 層続くものとして扱われる。DKS においては、DHT ノード数を N とすると、探索は $\log_k N$ ホップで行われ、個々の DHT ノードの経路表は $(k-1) \log_k N$ エントリとなる。本研究の DHT 実装は、このパラメータ k のアイデアを Chord に適用したものと見える。なお、もう 1 つの軸である、コストと信頼性のトレードオフについても、DKS ではパラメータ f を導入し、これに基づいてレプリカ制御を行っている。他に、アルゴリズムに応じた動的信頼性・静的信頼性について、Risson らの研究¹⁴⁾ にまとめられている。

トレーサビリティの研究としての DHT 応用の例には、Huang らの研究¹⁵⁾ がある。Huang らの研究では、EPC-IS を構成する各ノードによる DHT を作成し、これにより個品 ID から EPC-IS の URL を解決している。Huang らの研究は分散による規模拡張性の確保には触れているが、一方で、DHT 上のデータを管理する主体の責任とコスト負担が不明確になる問題と、トレーサビリティシステムが多様な商品カテゴリを取り扱う場合に商品カテゴリごとに異なる性質の個品 ID 空間を要求するケースに対応することが難しいという問題の 2 点については言及されていない。

3. 多様性および規模拡張性を実現する ID 解決機構の設計

3.1 要件と環境の定義

個品追跡トレーサビリティに用いる ID 解決機構の要件としては、以下の 3 つが存在する。

DNS を基本プロトコルとする：主として DNS が DNS ベースで構築されている、という理由から、クライアントからアクセスするためのプロトコルは DNS であることが望ましい。DNS クライアント (リゾルバ) はほぼすべてのインターネットノードで利用できるため、クライアント側は最小限の変更のみで個品 ID 解決機構が利用できる。

複数の異なる特性の ID 空間を統合する：個品を追跡するための ID 空間は、複数の商品カテゴリで必要とされている。ここで、個々の商品カテゴリそれぞれに独立した個品 ID 空間を構築すると、利用者は商品カテゴリごとに利用する個品 ID 解決手段を選択しなければならない。一方、個々の商品カテゴリはそれぞれ商品のライフサイクルや流通量、個品 ID 解決がもたらす付加価値とそれにもとない負担できるコストなどが異なり、単独の 1 つの ID 空間によって ID 解決手段を提供できるものではない。そこで、商品カテゴリごとに異なる特性の個品 ID 空間を用意し、これを 1 つのグローバルな商品 ID 空間に統合できる構成が求められる。

規模拡張性を持つ：個品追跡を行うためには、ある商品カテゴリに属する個品 ID すべてを含む ID を格納した個品 ID 空間を作り、また維持できなければならない。一方で、トレーサビリティシステムに当初から多大な投資は期待できないことから、小規模から円滑に拡張できる必要がある。また、季節性商品は 1 年の中で特定の時期に売上が集中するため、要求に応じた規模のサービスを動的に実現できる必要がある。

なお、本システムが想定する動作環境は次のとおり。本研究では、個品 ID 解決サービスを構成する各 DHT ノードは名前サービスのための専用ノードであり、一般に P2P 的な技術で想定されるような、クライアント PC によるものは想定しない。したがって、各 DHT ノードは、少数のデータセンタに集約された、十分に管理されたノードであるとする。その結果、ノードの故障発生頻度は通常のサーバと同程度 (MTBF は 1 年以上) であり、またノード間のネットワーク遅延は平均で 3 ミリ秒以下に抑えられるとする。なお、ネッ

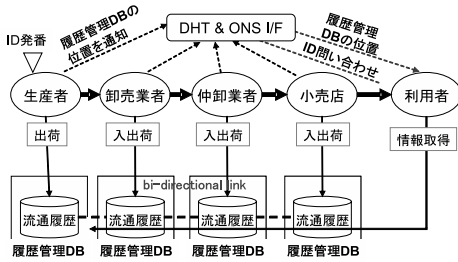


図 2 LOTR の全体像
Fig. 2 Overview of the LOTR.

トワーク遅延の仮定値は著者の環境から、東京都内に本拠があると思われる主要 ISP や企業などの WWW サーバまでの遅延を測定した値 (0.2~2.5 ミリ秒) に基づく。

3.2 システム構成

著者らが提案するトレーサビリティシステム (以後 LOTR: Linked Object TRaceability と呼ぶ) は、以下の要素からなる。図 2 にその概要を示す。

履歴管理 DB: 実体となる情報を管理するデータベース。LOTR における履歴管理 DB は流通経路の前後の「管理ドメイン」(生産・流通・卸・小売など、ある業者により物品の出入りが管理される領域) が持つ履歴管理 DB へのリンクを保持する機能が存在する。詳細は別文献 16) 参照。原則として 1 つの管理ドメインに対して代表する履歴管理 DB が 1 つ存在する。

ID-DB 解決: 以下の 3 つからなる、個品 ID からその ID に関連するデータベースの位置を発見するためのサービス。

ONS 互換 DNS: システムの入口として、ONS と類似した DNS サービスを持つ。これは、タグ ID から会社名レベル、あるいは商品カテゴリレベルの ID 解決の権限を、特定の DHT-DNS マウンタに移譲する。システム全体を通して 1 つ存在する。

DHT-DNS マウンタ: DNS の名前空間木に射影された商品 ID のうち「個品 ID」あるいは「商品カテゴリ ID+個品 ID」の部位すべてを、1 つまたは複数の DHT ノードに、DNS により権限を委譲 (NS RR¹⁷⁾) する。なお、どちらの構成になるかは、DNS の名前空間設計によって決まる (詳細は本稿の範囲外になるため省略)。商品 ID の生産者あるいは商品カテゴリに対して 1 つ存在する。

個品 ID 解決サービス (DHT): 個品 ID あるいは商品カテゴリ ID+個品 ID と、その ID に

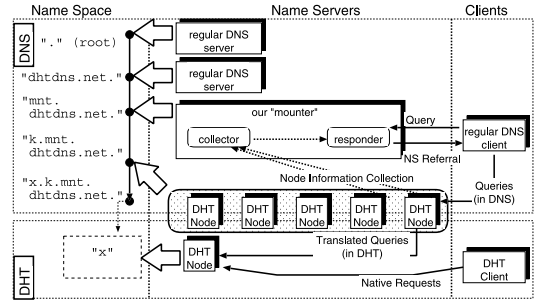


図 3 ID 解決に関する構成
Fig. 3 Designs on ID resolution.

対応するトレース DB の URL の対応付けを DHT の原理によって行う。各々の DHT ノードは DNS の問合せを受け付け、これを DHT の問合せに翻訳して解決し、DNS の応答 (NAPTR RR¹⁸⁾) を返す。

ONS 互換 DNS, DHT-DNS マウンタ, 個品 ID 解決サービスの 3 つの構成により、規模拡張可能かつ商品カテゴリごとに独立した性質を持たせられる個品レベル ID 解決が可能となる。

ID 解決に関連する構成例を図 3 に示す。図中左側が論理的な名前空間、中央がそれを構成するサーバ、右側が利用するクライアントである。また、図中上側が DNS, 下側が DHT により構成される。図中の DHT ノードはすべて、DNS における特定の名前空間 (この例の場合は k.mnt.dhtdns.net) に対する名前サーバとして振る舞う。多数の DHT ノードそれぞれに DNS client からの負荷を分散するために、DHT-DNS マウンタ (図中 mounter) は名前権限委譲のプロセスで、これら DHT ノードをラウンドロビンで利用する。

DHT-DNS マウンタ内には collector と responder の 2 つの機能が存在する。collector は DHT ノードの情報を動的に収集し、responder は DNS クライアントからの問合せに対して収集した DHT ノード情報を応答する。たとえば、x を個品 ID としたとき、x.k.mnt.dhtdns.net. の問合せがあった場合は、k.mnt.dhtdns.net. に対して権限を持つサーバとして収集した DHT ノード情報をクライアントに返す。このとき、DNS 名前空間におけるサーバの権限を示す NS レコードにはドメイン名しか記入できず、collector は IP アドレスしか収集していないため、それら IP アドレスから A/AAAA レコードを合成し、そのレコードの名前を NS レコードに格納する。

あるタグ ID に関連する情報を持つ管理ドメイン上の履歴管理 DB は、そのタグ ID を DNS 名前空間に射影した際に権限を持つ DNS 名前サーバを検索し、こ

れがDHTノードだった場合、自身のURLを登録する。各DHTノードにはDHTに格納されたエントリをDNSレコードに変換するルールが存在する。そのルールに沿った形式でURLを登録することで、DHTにより登録したデータをDNSから閲覧可能になる。なお、本DHTの実装では、DHTとの通信プロトコルにOpenDHT¹⁹⁾などと同等のXMLRPCと、独自プロトコルの2種が利用できる。

3.3 個品ID解決サービスの時間・コストトレードオフ制御

現時点では著者が持つ個品ID解決サービスに実装した方式はChordのみであるため、複数の方式による比較が難しい。そこで、実際に複数の方式を導入する代わりに、商品カテゴリが求める個品ID空間の性質のうち、ID解決にかかる時間と、メンテナンスのコストとのトレードオフに着目し、複数の性質の異なる個品ID空間を実現する。

具体的には、DKSと同様に、トレードオフを制御するパラメータを導入した。このパラメータにより、より多くの個品IDを導入する個品ID解決サービスはより多くの経路を持ち、より多くの定常トラフィック（隣接DHTの死活管理など）を発生させ、その結果より多くのコストを支払い、大規模なDHT網を一定時間以下で検索させることができるようになる。

Chordにおいては、探索のためのコストは、DHTノードの数を N としたときに

$$r = a \log_2 N + b \quad (1)$$

(ただし r ：再帰問合せ段数、 N ：DHTノード数、 a 、 b ：定数)といえる。

これに対し、アプリケーションが求めるDHTの特性に基づき任意に決定するパラメータ f_{amp} の導入と、 f_{amp} を活用するように、ChordアルゴリズムにおけるFingerの宛先を決定する`fix_fingers`関数の改良を行い、finger targetの選択の粒度を可変とする。特に、`fix_fingers`関数内で、元々のChordのアルゴリズムでは、 n をノード自身のハッシュ値空間上の位置、 m をハッシュ値空間のビット長として、

$$target[i] = n + 2^i \quad (0 \leq i < m)$$

とされていたtargetの選択を、

$$target[i] = n + 2^{\frac{i}{f_{amp}}} \quad (0 \leq i < m \cdot f_{amp})$$

とする。

図4に、計算により求めた、 $f_{amp} = 1.0, 2.0, 3.0, 4.0$ の場合のtargetの出現頻度の比較を示す。これは自ノードが0時の方向（図中上側）だった場合に、targetがどのように配置されるかを図示している。な

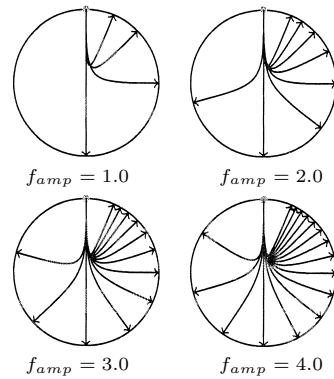


図4 $f_{amp} = 1.0, 2.0, 3.0, 4.0$ の場合のtargetの様子
Fig.4 Finger targets on $f_{amp} = 1.0, 2.0, 3.0, 4.0$.

お、Chordの空間を示す円のうち、 $\frac{1}{16}$ 相当未満に相当するtargetは図では省略している。一見して明らかかなように、targetの分布は f_{amp} を増大させるごとに密度が濃くなり、円周上のより多くのエリアに短いホップ数で到達できるようになる。反面、経路表維持のためには、それぞれのtargetについて死活管理を行う必要がある。これにより、定常トラフィックは f_{amp} に比例して増加する。

なお、本アルゴリズムはChordにパラメータを導入したものであるから、その基本的な特性を継承する。として、問合せに必要なホップ数は、 $O(\log N)$ 、隣接ノード数（経路数）および維持のためのトラフィック量は、 $O(f_{amp} \cdot \log N)$ となり、また、参加および離脱にかかるトラフィック量は、隣接ノードごとに問合せに必要なホップ数がかかり、 $O(f_{amp} \cdot \log^2 N)$ となると考えられる。

3.4 利用モデル

LOTRが想定する利用モデルを示す。

登場人物は、商品IDを生成する「生産者」、生産者を包含し、商品IDに対するトレーサビリティを提供する「業者」、商品IDに関連する情報を利用する「利用者（消費者を含む）」、個品ID解決サービスを提供する「個品ID解決サービス業者」である。

生産者は、生産する商品に紐付けられた商品IDがどのようなライフサイクルでどのような消費をされるかを助案し、「個品ID解決サービス業者」が提供するいくつかの選択肢から利用する個品ID解決サービスを選択し、契約する。生産者を含む、この商品を取り扱う業者（流通・加工・保管・卸・小売など）はこの個品ID解決サービスに対して、商品IDに関連する情報を持っていることを示すために、自身が持つ履歴管理DBのURLを登録する。

個品ID解決サービスは生産者と個品ID解決サー

ビス業者との関係（契約など）によって提供されるため、データ保持と責任の関係が明確になる。これは、P2P 技術の利用によって発生しがちなインセンティブ問題²⁰⁾が発生しにくい、という意味でもある。安定した個品 ID 解決サービスを提供することによりその商品に付加価値をつけられる生産者は、より高いコストを払ってより好ましい性質を提供する個品 ID 解決サービスを契約すればよい。

4. 評価

以下、(1) 問合せ量の総量と考えた際のシステム全体を通じてのボトルネックとなる responder のサービス構成の推定、(2) 保持すべきデータの総量に対する DHT ノード数およびその規模における問合せにかかる時間の推定、の 2 つにより LOTR を評価する。

4.1 ボトルネック

LOTR においては、処理可能な単位時間あたりの要求量を決定づける構成要素は、DHT-DNS マウンタである。個品 ID 解決サービス側の DHT ノードの数は要求量などに応じて随時増設可能だが、DHT-DNS マウンタの数には実質的な上限が存在する。これは、DNS では実用上メッセージを 1 つの分割されない UDP データグラムに収めることが求められていることによる。

DHT-DNS マウンタは、DHT ノードの情報を取得する機能 (collector) と、DNS 問合せに対してノード情報を NS RR の形で返答する機能 (responder) からなる。DHT-DNS マウンタ自体の性能は、responder 側の性能で一意に定まる。collector が十分な速度で動作しない場合は、同じ DHT ノードの情報を多数のクライアントに返答することで回避可能だからである。当然、その場合は一部の DHT ノードに問合せ負荷が集中する可能性がある。この問題は、十分な数の候補を並列に集めることで回避可能であると考え、本稿では考慮しない。

responder に対する負荷は、DNS の再帰リゾルバサーバによるものである。再帰リゾルバサーバは各個品にキャッシュを持つため、responder で指定された TTL (キャッシュ生存時間) の間はキャッシュレコードにより応答が行われる。したがって、TTL の調整により responder の実際の負荷は調整可能である。

4.1.1 responder の性能

表 1 に、試作した DHT-DNS マウンタの responder の性能 (1 秒間あたりの問合せ応答数) を示す。試作した DHT-DNS マウンタは 2 つあり、1 つは Python で書かれている多機能版、もう 1 つは C 言語で書かれ

表 1 DHT-DNS マウンタ responder の性能
Table 1 Performance of DHT-DNS Mounter's responder.

実装	qps 平均
多機能版	1,980
高性能版	7,970

ている高速版である。表 1 は、responder の性能限界を見積もるため、現状の実装における DNS 問合せ応答の速度 (qps: query per second) を ISC BIND に付属の queryperf により 10 秒間計測したものの 5 回の平均である。計測は、Dell PowerEdge 1750 (CPU: Intel Xeon 3.20 GHz x 2, Memory: 512 MB) 上で行った。なお、queryperf における並列化 (オプション-q) の影響は responder がシングルスレッドのため無視できる程度であった。

4.1.2 必要なサービス構成の推定

仮に、1 つのコンビニエンスストア (CVS) の POS で読み込まれる商品の種類の数は約 4,000 種^{*1}、日本国内の CVS 店舗数を 4 万店^{*2}とする。それぞれの店舗が独立した DNS キャッシュサーバを用いて個品 ID 解決をすると、最大で 1 億 6,000 万の (キャッシュサーバ、問合せ内容) の組合せが考えられる。

ここで、responder によるサービスの平均待ち時間を $h = \frac{1}{qps}$ とおく。また、クライアントは responder が返す TTL ごとに 1 回問合せを発行する、つまり、CVS 側では連続的に問合せが発生しており、キャッシュ生存時間が切れると必ず問合せが発生すると仮定する。responder がシングルスレッドであることから、これを $M/M/1$ の即時待ち行列と見なすと、Erlang の損失式

$$B = \frac{\lambda h}{\lambda h + 1}$$

により呼損率 B が求められる。

目標とする呼損率 B を 0.01 以下とする。これは、Jung らの測定²¹⁾における DNS の再送パケットの比率である 20~24% に比べて十分低い値として選択した。また、RFC1537²²⁾ で推奨されている一般的な TTL 値が 86,400 (1 日) であることと、DHT ノードが DNS と同等な水準で管理がされているという仮定から、TTL を 86,400 とおく。

この条件下で、目標とする呼損率 B を上まわらな

*1 <http://www.posbank.jp/> によると、商品点数 3,800 (2007 年 5 月時点)

*2 <http://www.d2.dion.ne.jp/%7EhMurata/conveni/tenpotable.html> インターネットタウンページに基づく調査結果 (2000 年)

表 2 目標とする呼損率 $B < 0.01$ を維持するための要求組合せ数 c の上限と DHT-DNS マウンタ数 S の下限Table 2 Upper bound of request combination c and Lower bound of the number of Mounters S .

	多機能版	高性能版
$c <$	1,728,001	6,955,682
$S \geq$	93	23

い範囲で 1 つの DHT-DNS マウンタが対応できる要求の組合せ数 c と、日本の CVS 全体 (1 億 6,000 万組合せ) に対応するために必要な DHT-DNS マウンタの数 S を、表 2 に示す。

DNS のルートネームサーバ同等の並列度を 13 を DHT-DNS マウンタがサービスする各ドメインに適用することを考えても、すべての製品を 1 つのドメインで支えることは困難であることが予想できる。一方で、図 1 に示したように、商品 ID には構造が存在し、DHT-DNS マウンタあるいは DNS はこの構造に沿ってそれぞれの商品の名前空間を分離できる。したがって、これを利用することで 2 (高性能版: $2 > 23/13$) ないし 8 (多機能版: $8 > 93/13$) 以上に名前空間を分割することにより、現状の DHT-DNS マウンタにより構成する名前サービスで、仮定した要求量に答えられるといえる。3.4 節で述べた、商品の製造会社別に分割するモデルであれば問題なく答えられる要求量といえる。

4.2 DHT 規模と問合せにかかる時間の推定

本研究では、3.1 節に示したように、100 億個の電子タグ ID を検索するとき検索に利用できる時間を平均 0.3 秒とし、この条件を満たす構成の実現可能性を検討する。

検討は、必要なノード数の推定、そのノード数における DHT 中の再帰問合せ段数の推定を行い、最終的に 100 億個の ID を検索するための必要時間の推定値を算出する。

4.2.1 必要ノード数の推定

必要ノード数はノードローカルに存在するデータテーブルの適切なサイズを求め、テーブルの総量をこれで割ることによって行う。

ノードローカルなテーブルの管理には、Berkeley DB の実装を用いた。4.1.1 項と同等の Linux OS 環境において大量のデータを格納したテーブルを生成したところ、 10^6 規模までのデータは問題なく取り扱えるが、 10^7 規模のデータのテーブルにおいては、データの書き込み性能が劣化した。表 3 に、本研究で用いているテーブル管理ルーチンを経由してデータの書き込みを行うテストプログラムの実行結果を示す。

表 3 テーブル読み書き時間の平均 (ミリ秒)

Table 3 Table read/Write time average (milliseconds).

テーブル規模	書き込み	読み出し
10^5	0.15	0.13
10^6	0.55	0.12
10^7	3.22	0.44

ここで、ノード 1 台あたり平均 10^6 個の ID を管理するとして、 10^{10} 個の ID を管理するには 10^4 台のノードが必要であるといえる。

4.2.2 再帰段数の推定

文献 6) では、試作した DHT で採用している基本アルゴリズム (Chord) の再帰問合せの深さは、 $0.5 \log_2 N$ であるとされている。ここで単純かつ理想的な二分木探索が行われているとすると、 10^4 台のノードにおいては 6~7 段程度 ($0.5 \log_2 10^4 \approx 6.64$) の再帰が行われると予測できる。しかし、実装したアルゴリズムはテーブル空間上の近傍情報の処理に最適化が施されていたり、逆に確実に対象ノードに到達するために安全な経路を取り余分なホップ数をかけたり、などといった理由により、理想的な二分木探索とは異なる可能性がある。そこで、計測が容易で計算機の性能などの条件に左右されない「再帰問合せ段数」の平均値のモデルを 3.3 節の式 (1) とし、パラメータ a と b を実測により求めることとした。

DHT 単独の規模拡張性計測において、次の 2 点を目的とした

- DHT 自身が式 (1) をモデルとし規模拡張可能なことを確認する。
- 検索の細かさを規定するパラメータ f_{amp} を変更し、それぞれにおいて式 (1) の定数 a, b を求める。

計測には、Pentium III 1 GHz 主記憶 512 MB の計算機が 208 台存在する StarBED クラスタ A を利用した。StarBED は、NICT 北陸リサーチセンターに存在する、PC の実機を用いた分散エミュレーション環境であり、大規模実験制御機能²³⁾ が提供されている。

これらの物理ノードに、OS として Linux 2.4.27 を PXE Boot によりロードし、DHT 環境を実現した。また、実際にはより多数の DHT ノード環境を模擬するため、1 物理ノードあたり 20DHT ノードを動作させた。この物理ノード数を n とおく。

この n 台の物理ノードとは別に、もう 1 つの物理ノードを専用に確保して、DHT の起動時に他のノードへの連絡を行うプロセス (seeding node) を 1 つ起動している。これは起動後は通常の DHT ノードとして振る舞うため、合計 DHT ノード数 N は $N = 20n + 1$ となる。

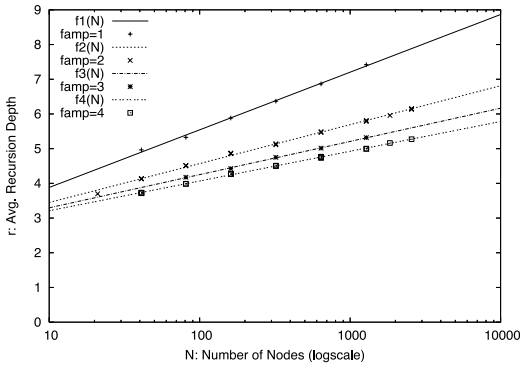


図5 パラメータ f_{amp} の効果
Fig. 5 Effect of f_{amp} .

表4 f1 から f4 までのパラメータ a , b の具体値と, 10,000 ノード時の予測再帰段数 r および問合せ応答予測時間 Q
Table 4 Estimated values of parameter a , b from f1 to f4, Estimated recursion depth r , and estimated response time Q .

	a	b	r	Q
f1	0.500	2.22	f1(10,000)	8.86 292
f2	0.338	2.33	f2(10,000)	6.82 225
f3	0.288	2.34	f3(10,000)	6.17 204
f4	0.258	2.35	f4(10,000)	5.78 191

実験では, パラメータ f_{amp} が 1, 2, 3, 4 の 4 通りの場合について計測した. 結果を図 5 に示す.

横軸は DHT ノード数 (21 から 2561), 縦軸が再帰の深さである. 試行回数は全体で 48 回であり, 一部のパラメータ組のみ複数回試行し, 誤差が十分少ないことを確認している. $f_{amp}=1, 2, 3, 4$ はそれぞれの実験結果であり, それぞれに対応する f1, f2, f3, f4 は回帰分析により式 (1) を近似した結果である. 式 (1) に対応する回帰分析結果のパラメータを表 4 左に示す. なお, f1 の場合は Finger の構成方法は Chord と同等であり, $a = 0.5$ であるが, 実装上経路を最適化しない部分から定数 b が発生し, 約 9 段の再帰段数となっている.

4.2.3 100 億個 ID の問合せ時間の推定

前出の規模拡張性の検討結果と, 4.2.1 項の検討から, $f1(10,000)$ から $f4(10,000)$ を推測し, その結果を表 4 右に示す. なお, 図 5 に見てとれるように, 4.2.2 項で推測したパラメータは十分に事象を表していると考えられる. ここでの誤差は最大で約 4.7% であり, 2 つのケースを除いたすべてのパラメータが誤差 2.5% 未満で実測値に一致した.

これらから, 本研究で開発した DHT の実装を用いた, 10,000 ノード DHT の問合せ再帰問合せ段数の平均値は, 5.8 ~ 8.9 程度であると推測できる.

1 つの問合せが応答されるまでの DHT 内の処理にかかる時間 Q は次式で表せる.

$$Q = (t_n + t_r) \times h + t_l$$

(ただし, t_n : ノード間のネットワーク遅延, t_r : ノード再帰問合せ応答時間, h : 再帰問合せ段数, t_l : 最終段テーブル検索時間)

1 段の再帰にかかる平均的な時間は, DHT1 ノードの再帰問合せ応答 (メッセージ処理および再帰経路のルックアップ) にかかる時間と同等である. 実測で 30 ミリ秒程度という数値を得ている. 3.1 節で述べたとおり, DHT ノード間の遅延を 3 ミリ秒以下に抑えたとする. なお, t_l は表 3 から, 10^6 のテーブル規模であれば 1 ミリ秒未満なので無視できる.

これらの実測値および仮定から, およその Q (ミリ秒) が

$$191 < Q < 292$$

程度であることが推定できる.

5. 考 察

3.1 節で本 ID 解決機構の 3 つの要件を述べた. まず, DNS を基本プロトコルとする要件は, DHT 各ノードに DHT レコードと DNS レコードの変換機構を備えることと, DHT-DNS マウンタによって DHT ノードに対して権限委譲を行うことによって満たした.

多様な特性の個品 ID 空間の実現については, 3.3 節で原理の 1 つを示し, 3.4 節で利用モデルを示し, 4.2 節の中で, f_{amp} の利用による問合せ応答速度の調節について検証した. 図 5 で, f_{amp} のパラメータ変更により, 特性の 1 つ, 環境維持のコストと問合せ解決の時間のトレードオフを選択可能なことを示した. f_{amp} およびノード数の異なる複数の DHT を利用し, 商品カテゴリごとに適切な DHT を利用することで, 問合せ応答の速度とコストのバランスを要求に応じて選択可能な環境が実現できる.

規模拡張性については, DHT-DNS マウンタ部分については 4.1.1 項で, DHT 部分については 4.2 節で評価した. DHT-DNS マウンタの性能がトレーサビリティシステムの一応用に十分な性能を持っていることを示し, また, DHT 自体の問合せにかかる時間に密接に関係する再帰段数を実験から具体的に求めた. 個品 ID 解決にかかる問合せ応答に対する平均値が数百ミリ秒オーダーなので, 実際の DB への問合せや, ADSL など接続されたクライアントサイトまでの伝送遅延まで含めると 1 秒程度の時間が必要である. ユーザとの対話型システムとしては許容範囲だが, ロット単位の検品などといったバルク処理を行う場合は, 商品

単位に同期的に処理を行っていたのでは商品数に比例した時間がかかるため、ID 解決および検品にともなうデータ登録処理などの非同期並列処理化が必須であろう。

本稿では DHT が 1 つまたは近傍少数のデータセンタに格納されているという仮定において分析を行った。もし、ネットワーク遅延や信頼性の実態が仮定と異なる場合は、パラメータの調整や他の機構の導入などを行うことにより、要求される問合せ応答時間内で処理を行わせるよう努力することとなる。さらに、本研究の自然な延長として、DNS 応答部分を付加した他の DHT アルゴリズムや実装を導入することで、個品 ID 空間の選択肢を広げることが考えられる。たとえば、Pastry²⁴⁾ を利用することで、より短いホップ数で検索を行ったり、あるいは多数の拠点にノードを地理的に広域に分散した場合に、近傍性を考慮した経路選択を行ったりすることにより、伝送遅延の効果を最小限にとどめることなどが考えられる。

6. 結論と今後の課題

本研究では、個品追跡可能なトレーサビリティシステムを実現する基礎として、DHT と DNS を併用したハイブリッド ID 解決機構についての検証を行った。100 億 (10^{10}) 個の商品 ID を検索するために、10,000 ノードを想定し、これらの中で対象となるデータの探索を行うことでどの程度の時間がかかるかを推定した。約 2,500 ノードまでの規模の実験と、文献に基づくモデルを適用し、10,000 ノードの再帰段数を推測したうえで、実測した問合せ応答の 1 段あたりの処理時間などから、具体的な数値を求めた。10,000 ノードから構成される DHT は、平均して 6 段～9 段の再帰問合せを行う。その結果平均 179～275 ms. 程度の時間を ID 解決で消費することを推定した。

本研究により、10 の 10 乗規模の個品トレーサビリティシステムが実現可能であることを示した。ただし、実用に至るためには、ONS をはじめとした既存の機構との共存や、DHT においたデータの保護やバックアップなどといった機構が必要である。さらに、より適切な DHT 方式や実装の選択肢を作るために、DHT-DNS 変換などといった個品 ID 解決サービス独自の作り込み部分のモジュール化による他の実装への移植や、より適切なアルゴリズムの取り込みなども必要である。今後は DHT を利用した個品トレーサビリティシステムを実現するために、実用化の障害となる問題について明らかにしたうえで、DHT の健全性の維持をはじめとした研究を行う予定である。

謝辞 本研究の一部は、総務省による平成 18 年度「電子タグ高度利活用に関する研究開発」に係る委託研究の成果によるものである。本技術のフィールド実証については、日本電気(株)、横河電気(株)、エヌ・ティ・ティ・コミュニケーションズ(株)の共同研究者との協力によって行った。また、大規模な DHT 実験環境の構築については、ICT 研究開発機能連携推進会議、NICT 北陸リサーチセンター宮地氏、北陸先端科学技術大学院大学情報科学研究科知念氏の支援を受けた。ここに感謝の意を表する。

参考文献

- 1) Traub, K., Allgair, G., Barthel, H., Burstein, L., Garrett, J., Horgan, B., Rodrigues, B., Sarma, S., Schmidt, J., Schramek, C., Stewart, R. and Suen, K.: The EPCglobal Architecture Framework, EPCglobal Standard Specification (2005).
- 2) EPCglobal: EPC Tag Data Standards Version 1.3, EPCglobal Ratified Specification (2006).
- 3) Wessel, R.: Austrian Bank Finds RFID Yields Big Returns When Tracking Loan Files (2007). <http://www.rfidjournal.com/article/view/3160/>
- 4) NTT データ株式会社: 青果物トレーサビリティの実証実験を開始, プレスリリース (2004). <http://www.nttdata.co.jp/release/2004/012800.html>
- 5) T-Engine フォーラム: 農林水産省平成 17 年度ユビキタス食の安全・安心開発事業『ユビキタス食品情報基盤システム実証事業』に関する記者会見および公開説明会, 見学会のご案内: 「ユビキタス食品情報基盤システム実証実験」説明会, プレスリリース (2006). <http://www.t-engine.org/news/pdf/TEP060215-u04.pdf>
- 6) Stoica, I., Morris, R., Karger, D., Kaashoek, M.F. and Balakrishnan, H.: Chord: A scalable peer-to-peer lookup service for internet applications, *Proc. ACM SIGCOMM* (2001).
- 7) 尾崎 哲, 土井裕介, 若山史郎: 小規模から円滑に拡張できる商品トレーサビリティシステム, 東芝レビュー, Vol.60, No.8, pp.27-31 (2005).
- 8) Doi, Y.: DNS meets DHT: Treating Massive ID Resolution Using DNS Over DHT, *Proc. SAINT 2005*, pp.9-15 (2005).
- 9) 高玉広和, 矢島健一, 水越康博, 竹内章平: 電子タグを活用した食肉トレーサビリティ実証実験, 電子情報通信学会ソサイエティ大会講演論文集, Vol.2006, No.1, pp.S-53-S-54 (2006).
- 10) IT 戦略本部: e-Japan 重点計画 - 2004 (2004).
- 11) EPCglobal: Object Naming Service (ONS) version 1.0, EPCglobal Ratified Specification (2005).

- 12) Cox, R., Muthitacharoen, A. and Morris, R.: Serving DNS using a Peer-to-Peer Lookup Service, *Proc. 1st International Workshop on Peer-to-Peer Systems (IPTPS)*, Cambridge, MA (2002).
- 13) Alima, L.O., El-Ansary, S., Brand, P. and Haridi, S.: DKS(N, k, f) A family of Low-Communication, Scalable and Fault-tolerant Infrastructures for P2P applications, *The 3rd International workshop on Global and P2P Computing on Large Scale Distributed Systems (CCGRID 2003)* (2003).
- 14) Risson, J. and Moors, T.: Survey of Research towards Robust Peer-to-Peer Networks: Search Methods, *Computer Networks*, Vol.50, No.17, pp.3485–3521 (2006).
- 15) Huang, D., Verma, M., Ramachandran, A. and Zhou, Z.: A Distributed ePedigree Architecture, *Proc. FTDCS 2007* (2007).
- 16) 若山 史郎, 土井裕介, 井上 淳: TrackBack 手法に基づくトレーサビリティシステムの設計と実装, 電子情報通信学会技術研究報告 (NS2004-321), pp.379–382 (2005).
- 17) Mockapetris, P.: DOMAIN NAMES – CONCEPTS AND FACILITIES, IETF RFC 1034 (1987).
- 18) Mealling, M. and Daniel, R.: The Naming Authority Pointer (NAPTR) DNS Resource Record, IETF RFC 2915 (2000).
- 19) Rhea, S., Godfrey, B., Karp, B., Kubiatowicz, J., Ratnasamy, S., Shenker, S., Stoica, I. and Yu, H.: OpenDHT: A Public DHT Service and Its Uses, *Proc. ACM SIGCOMM 2005* (2005).
- 20) Adar, E. and Huberman, B.: Free riding on gnutella, *First Monday*, Vol.5, No.10 (2000).
- 21) Jung, J., Sit, E., Balakrishnan, H. and Morris, R.: DNS Performance and the Effectiveness of Caching, *Proc. ACM SIGCOMM Internet Measurement Workshop '01*, San Francisco, California (2001).
- 22) Beertema, P.: Common DNS Data File Configuration Errors, IETF RFC 1537 (1993).
- 23) Miyachi, T., Chinen, K. and Shinoda, Y.: StarBED and SpringOS: Large-scale General Purpose Network Testbed and Supporting Software, *International Conference on Performance Evaluation Methodologies and Tools (Valuetools) 2006* (2006).
- 24) Rowstron, A. and Druschel, P.: Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems, *Proc. 18th IFIP/ACM International Conference on Distributed Systems Platforms (Middleware 2001)*

(2001).

(平成 19 年 6 月 7 日受付)

(平成 19 年 12 月 4 日採録)



土井 裕介 (正会員)

平成 12 年慶應義塾大学大学院政策・メディア研究科修士課程修了。現在 (株) 東芝研究開発センター通信プラットホームラボラトリーで分散システム関連研究に従事。電子情報通信学会, ACM 各会員。



若山 史郎

平成 15 年慶應義塾大学大学院政策・メディア研究科修士課程修了。現在 (株) 東芝研究開発センター通信プラットホームラボラトリーで RFID 応用技術, ユビキタスネットワークワーキング関連研究に従事。



石山 政浩

平成 6 年横浜国立大学大学院工学研究科修士課程修了。(株) 東芝研究開発センター通信プラットホームラボラトリー研究主務, 工学博士。ネットワークレイヤを中心としたモバイルコンピューティングおよびセキュリティ技術の研究開発に従事。



尾崎 哲

平成 7 年横浜国立大学大学院工学研究科博士課程修了 (株) 東芝研究開発センター通信プラットホームラボラトリー研究主務, 工学博士。ユビキタスネットワークワーキング, ネットワーク家電, ネットワークセキュリティ関連研究開発に従事。IEEE, 電子情報通信学会会員。



井上 淳 (正会員)

昭和 60 年東京大学工学部計数工学科卒業。(株) 東芝イノベーション推進部経営変革エキスパート。インターネットプロトコル全般, 特に IPv6, IP モビリティ, ネットワークセキュリティ関連の研究開発に従事。ACM 会員。