

複数属性を持つデータの属性値範囲指定検索用分散ネットワーク構成方式の提案

松田 哲史

三菱電機株式会社 情報技術総合研究所

〒247-8501 神奈川県鎌倉市大船 5-1-1

E-mail: Matsuda.Tetsushi@dh.MitsubishiElectric.co.jp

要約 複数の属性を持つデータに対して、属性値の範囲を条件に指定してデータ検索を行うことを考える。こうした範囲指定検索を、多数の装置から構成される分散データ管理ネットワークで実現するための、ネットワーク構成方式を提案する。

A Proposal of distributed data management network for searching data with
multiple attributes by specifying attribute value ranges

Tetsushi Matsuda

Information Technology R&D Center, Mitsubishi Electric Corporation

5-1-1, Ofuna, Kamakura-shi, Kanagawa, 〒247-8501

E-mail: Matsuda.Tetsushi@dh.MitsubishiElectric.co.jp

Abstract Our objective is to enable searching datum with multiple attributes by specifying the ranges of multiple attributes in a distributed data management network. We propose a method to build such a network.

1. はじめに

近年、Peer to Peer (P2P) ネットワークによるファイル交換サービスが普及してきているが、ファイル交換でユーザが目的のファイルを検索する時に、正確なタイトルや作者名が分かっていることも多い。また、センサーネットワークからの情報取得において、センサーやデータの属性が満たす条件（例 位置=新宿近辺 AND センサー種別=温度センサー）を指定して、目的データを取得することが必要になると考えられる。こうした応用に P2P ネットワーク技術を適用するために、検索対象データの属性情報（例 タイトル, 作者, 作成年度）に対する検索条件として、属性値の範囲指定を可能とするデータ検索方式が必要になると考えられる。

構造的 P2P ネットワークを構築するために広く用いられる分散ハッシュテーブル(DHT)技術は、検索キーに対する完全一致検索を元々想定しており、こうした応用で必要となる範囲指定検索をサポートするためには拡張が必要となる。本論文では、検索対象データの複数属性に対する範囲指定を可能とする検索方式について、検討中の内容を報告する。

2. 関連する研究

データの属性値の範囲条件を指定して検索を行うための分散データ管理ネットワーク構成方式に関する、既存の研究([1], [2], [6])の概要と、どの様な問題点があるかを説明する。

・Mercury[1]

データが持つ各属性に対して属性ハブと呼ぶリング形状の分散データ管理ネットワークを構成する。属性ハブを構成する各ノードは、その属性値のある範囲に属するデータを管理しており、ノードは範囲の大小関係の順序で属性ハブ上に並ぶ(図 1)。属性ハブ内でのデータ登録/削除要求とデータ検索要求転送は、隣接ノードへのリンクと、調和分布に従い確率的に決まるノード数だけ離れた位置にある k 個 ($k=0(1)$) のノードとの間に設定される長距離リンクを用いて行う。これにより、 $O(\log(n)/k)$ (n =属性ハブを構成するノード数) ホップで、任意の値を管理するノードへのルーティングを行うことが可能となる。よって、属性値を指定した検索要求は $O(\log(n)/k)$ ホップで処理可能となる。属性値の範囲を指定した検索要求の転送に必要なホップ数は、検索条件に指定される属性値の範囲の分布に依存して、 $O(\log(n)/k)$ から $O(n)$ ホップとなる。

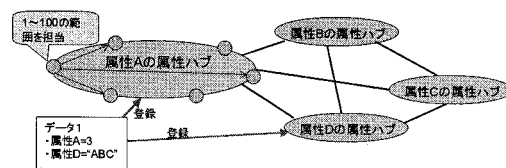


図1 Mercuryの属性ハブ構成

Mercury の特徴は、属性ハブを構成する個々のノードが管理する属性値の範囲のサイズを、ノードの管理データ数や検索要求処理数の負荷状況に応じて変化させることで、動的負荷分散を行うことにある。動的負荷分散に必要な、属性ハブを構成するノードの負荷と属性値の範囲のサイズを計測するために、ランダムサンプリングを用いる。

Mercury では、あるデータが複数の属性を持つ場合に、各属性に対する属性ハブ全てに、データの全属性の情報を登録する必要がある。このため、データが持つ属性数が増えると、1 つのデータの全属性値の情報のコピーを属性数個分ネットワークで管理する必要があるため、データ格納領域の無駄が多くなるという問題がある。また、全ての属性ハブ間にフルメッシュでリンクをはる必要があるため、属性数が増えると属性ハブ間のリンク数が属性数の 2 乗に比例して増えるという問題点もある。検索要求は、検索要求に指定される条件に含まれる属性の中のいずれか 1 つを選び、その属性に対する属性ハブへ転送し、その属性ハブ内で検索を行うことで処理を行う。このため、検索要求に複数の属性の範囲条件が指定されている場合でも、検索要求の転送先となるノードを、全ての属性条件を利用して絞り込むことができないという問題もある。

・SkipNet [6]

Skip List と呼ばれるデータ構造を、ネットワークポロジ維持のための処理負荷が大きくなりないように、近似的に実現する分散データ管理ネットワーク構成方式である。SkipNet を構成するノードには、Name ID と Numeric ID の 2 つの ID が割当てられ、ノードは、Name ID の大小関係の順にリング状に配置され、指定された Name ID/Numeric ID のデータを管理するノードの検索を、高い確率で $O(\log(n))$ (n =SkipNet を構成するノード数) ホップで可能とする。データの 1 つの属性値を Name ID として SkipNet にデータを登録することで、1 つの属性値に対する範囲検索を実現することが可能である。SkipNet では、属性値を指定した検索要求を、検索要求を $O(\log(n))$ ホップ転送することで処理可能とする。属性値の範囲を指定した検索要求の転送に必要なホップ数は、検索条件に指定される属性値の範囲の分布に依存して、 $O(\log(n))$ から $O(n)$ ホップとなる。

SkipNet には、複数属性の範囲検索が対象外であるという問題と、負荷分散を行う場合に Name ID の一部をハッシュ値計算用データとして用いたハッシ

ュによる負荷分散を行っており、Name ID のその部分を範囲検索に利用することができなくなってしまうため、負荷分散と範囲検索を両立させることが困難であるという問題がある[1]。

3. 提案方式

3. 1 提案方式のネットワーク構成

2 でみた既存の方式の中で Mercury が優れていると考えられるが、Mercury における問題点を解決するために、以下の性質を持つ分散データ管理ネットワークを構成する。

- (a) 属性ハブ間にフルメッシュをはることなく属性ハブ間で通信することが可能。
- (b) 複数属性を持つデータでも、分散データ管理ネットワークに登録する属性情報を 1 つとする。
また、検索要求に指定される属性値の範囲条件を全て使用して、検索要求の転送先となるノードを絞り込むことを可能とする。
- (c) 管理データ数や検索要求処理数に応じた負荷分散を可能とする。
- (d) 属性値を指定した検索要求の転送ホップ数と、属性値の範囲を指定した検索要求の転送ホップ数を、Mercury と同程度に抑えられる。

このために、以下の 3 種類のネットワークから分散データ管理ネットワークを構成する。以下では、検索対象データの属性数の上限値が存在すると仮定し、その上限値を d とする。各属性 A_i ($1 \leq i \leq d$) の値の範囲は、 $-R_i$ より大きく R_i 以下とする。 $-R_i$ は、属性 A_i の値が不明であることを表す値として使用する。“次元の呪い”の問題があるため、提案方式でも、データ管理ネットワークを構成するノード数 n に対して d の値が大きい場合には、効率的な検索を実現することは困難であることには注意が必要である。

(1) データネットワーク

データ実体と属性値情報を管理するために、属性を座標軸とする d 次元空間を超直方体に分割して管理するネットワーク。

データネットワークを構成するノードは、 d 次元空間の超直方体で表されるエリアを管理し、管理する超直方体のエリアに属するデータを管理する。

データネットワークは、CAN[4]と同様な構成をとるが、CAN の場合はハッシュ値を用いて負荷分散を実現するのにに対して、データネットワークでは(3)の負荷分散ネットワークを用いて動

的に負荷分散を行う点が異なる。

データネットワーク内で、隣接超直方体を管理するノード間で検索要求を転送することで、属性値の範囲指定検索を実現する。

(2) インデックスネットワーク

データネットワークを構成する各ノードが管理する超直方体のエリア情報を管理するネットワーク。検索条件に指定される属性値の範囲条件から、最初に検索要求を転送する先となるデータネットワークのノードを決定するために使用する。登録データ数は、データネットワークを構成するノード数に等しい。

インデックスネットワークは、各属性に対して1つ作ることが可能。属性Xに対するインデックスネットワークはMercuryの属性ハブの考え方で構成する。すなわち、インデックスネットワークを構成するノード1は属性Xのある範囲を管理しており、d次元超直方体2の属性Xの値範囲が、ノード1が管理する属性Xの値範囲と重なりを持つ場合、ノード1は超直方体2のエリア情報を管理する。

(3) 負荷分散ネットワーク

データネットワークを構成する各ノードの処理負荷の情報を集約し、ノード数とノードの処理負荷の最大値と最小値の情報を集め、データネットワークでの負荷分散処理を制御するためのネットワーク。

負荷分散ネットワークは、データネットワークを構成するノードでChord[3]のネットワークを構成し、Chordのネットワーク上に論理ツリーを構成することで実現する。負荷分散ネットワーク上での付加情報集約方式は、RanSubアルゴリズム[5]に基づく。RanSubアルゴリズムで使用するツリーとして、Chordのネットワーク上にFinger Tableを用いて構成する論理ツリーを用いる。

論理ツリーのを用いたRanSubアルゴリズムの実行方法は以下の通り。

ChordネットワークでのノードIDで最も小さな値を持つノードをルートノードとする。

ルートノードは、自ノードのFinger Table (Chordで使用する他ノードへのリンクテーブル)に登録されるノードに対して、RanSubでの情報問い合わせメッセージを送信する。情報問い合わせメッセージには、宛先ノードのFinger Tableでのインデックス値(Finger Tableインデ

ックスと呼ぶ)を示す。

情報問い合わせメッセージを受信したノードは、自ノードのFinger Tableに登録されるノードの中で、メッセージに示されるFinger Tableインデックスよりも小さな値のインデックスを持つノードに対して情報問い合わせメッセージを送信する。この時、情報問い合わせメッセージに示すFinger Tableインデックス値は、自ノードのFinger Tableを元に更新する。

以上により、論理ツリーに沿って、ルートノードから子ノードへと情報問い合わせメッセージが転送される。情報問い合わせメッセージに対する応答は、情報問い合わせメッセージの送信元ノードに返す。

3. 2 提案方式でのデータ検索・登録・負荷分散処理

上記の3つのネットワークを用いて、データの検索処理、登録処理、負荷分散処理をどの様に行うかを説明する。

3. 2. 1 データ検索処理

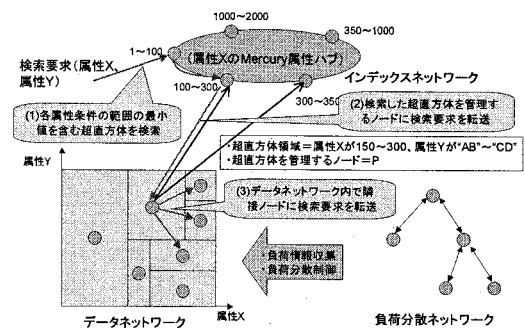


図2 データ検索方式

(0) 検索要求元の端末1は、検索要求に指定される属性の中の任意の属性を選び、その属性に対応するインデックスネットワークの任意のノードAに検索要求を送信する。

(1) 検索要求を受信したノードAは、検索要求に指定される属性値範囲の下限(未指定の属性 A_i については、 $-R_i$ [=不明]が指定されたと見なす)となる点を含む超直方体を管理するデータネットワークのノードBを、Mercuryの属性ハブ上での検索方式に従って検索する。

検索条件に指定される属性値範囲が $L_i \leq A_i \leq U_i$ ($1 \leq i \leq d$) の場合、属性値範囲の下限を、d

次元空間の点($L1, \dots, Ld$)と定義する。

- (2) ノードAは、検索要求をノードBへ転送する。
- (3) 検索要求を受信したノードBは、自ノードが管理するデータで検索条件を満たすものを求め、その情報を端末1へ送信する。

また、各属性 A_i について、検索条件で指定される属性 A_i の範囲が、自ノードが管理する超直方体の属性 A_i の最大値よりも大きな値を含む場合は、属性 A_i 軸上でより大きな値のエリアを管理する隣接超直方体を管理する、データネットワークのノードC(超直方体の分割方法によっては複数の場合あり)へ検索要求を転送する。

ノードCが検索要求を重複受信したことを識別できる様に、ノードBが検索要求を転送する際に、(ノードBの識別子, 検索要求ID)を付加する。

- (4) 検索要求を受信したノードCは、受信した検索要求に付与される(ノードBの識別子, 検索要求ID)の値が受信済みでないかをチェックし、受信済みでなければ自ノードが管理するデータで検索条件を満たすものを求め、その情報を端末1へ送信し、(ノードBの識別子, 検索要求ID)の情報を記録する。記録した(ノードBの識別子, 検索要求ID)の情報は、一定時間(=1つの検索要求が重複して転送される可能性がある時間)後に削除する。受信済みの場合は、何もしない。

また、検索要求を属性 A_j 軸上の隣接超直方体を管理するノードから受信した場合、属性 A_i ($j \leq i$)となる各属性 A_i について、検索条件で指定される属性 A_i の範囲が、自ノードが管理する超直方体の属性 A_i の最大値よりも大きな値を含む場合は、属性 A_i 軸上でより大きな値のエリアを管理する隣接超直方体を管理する、データネットワークのノード(超直方体の分割方法によっては複数の場合あり)へ検索要求を転送する。

3. 2. 2 データ登録処理

- (0) 登録要求元の端末1は、登録要求に指定されるデータの属性の中の任意の属性を選び、その属性に対応するインデックスネットワークの任意のノードAに登録要求を送信する。
- (1) 登録要求を受信したノードAは、登録要求に指定される属性値(未指定の属性 A_i については、 $-R_i$ [=不明]が指定されたと見なす)を表す点を含む超直方体を管理するデータネットワークのノードBを、Mercuryの属性ハブ上での検索方

式に従って検索する。

- (2) ノードAは、登録要求をノードBへ転送する。
- (3) 登録要求を受信したノードBは、自ノードが管理するデータに、登録要求で指定されるデータを追加する。

3. 2. 3 負荷分散処理

- (0) 負荷分散ネットワーク上で、RanSubアルゴリズムを定期的に行うことにより、論理ツリーのルートノードが、ノード数、最大負荷ノードが何かとその処理負荷、最小負荷ノードが何かとその処理負荷、平均負荷の情報を求める。

- (1) 最大負荷が平均負荷の α 倍より大きく、かつ、最小負荷が平均負荷の $1/\alpha$ より小さい($\alpha > \sqrt{2}$ を満たす定数)となっている場合、論理ツリーのルートノードが負荷分散処理が必要と判断する。
- (2) 負荷分散処理が必要な場合、最小負荷ノードに対して、データネットワークから一旦離脱し、最大負荷ノードを接続ノードとしてデータネットワークに再度参加する様に指示を行う。
これにより、最小負荷ノードが最大負荷ノードの負荷を分担するようになる。最小負荷ノードが管理していたデータは、最小負荷ノードが管理する超直方体の隣接超直方体を管理するノードのいずれかが引継いで管理する。

4. 定性的評価

提案方式の定性的評価を行う。データネットワークを構成するノード数(=負荷分散ネットワークを構成するノード数)を n とする。インデックスネットワークは、データネットワークを構成する全ノードから構成する必要は無く、 d 個ある属性各々に対して n/d 個のノードでネットワークを構成すれば良いが、以下では n 個のノードから構成するとする。

4. 1 負荷分散処理に要するメッセージ数

負荷分散処理に必要な情報を収集するために、負荷分散ネットワーク上で送信するメッセージ数を検討する。

1回の情報収集で、各ノードは、ツリーの親ノードから1メッセージを受信し、親ノードへ1メッセージを送信するので、送信するメッセージ数は、 $2n$ となる。

各ノードが送信するメッセージの数の最大値は、そのノードのChordのFinger Tableに登録される異なるノードの数となるので、高い確率で $O(\log(n))$

となる。

4. 2 データ検索/登録/削除処理のメッセージホップ数

データ検索/登録/削除処理において、インデックスネットワークを検索してデータネットワークのノードに要求を転送するところまでの処理は、Mercuryの属性ハブにおける検索要求処理を用いているので、 $O(\log(n)/k)$ (k は、インデックスネットワークで各ノードがはるロングリンク数) ホップとなる。

データ登録/削除要求の場合、インデックスネットワークでの検索で得られたデータネットワークのノードが処理を行うため、必要なホップ数は $O(\log(n)/k)$ となる。

検索要求が属性値を指定する検索である場合は、インデックスネットワークでの検索で得られたデータネットワークのノードが処理を行い、データネットワーク内での検索要求の転送は行われないため、必要なホップ数は $O(\log(n)/k)$ となる。

属性値の範囲を指定する検索要求の場合の、データネットワーク内での検索要求の転送ホップ数については、以下の様に考えられる。データネットワークは負荷分散を行っており、処理負荷が高い領域に処理負荷に比例した数のノードを配置することで、処理負荷を α 倍から $1/\alpha$ 倍の範囲で均一化している。よって、属性 A_i ($1 \leq i \leq d$) が作る $-R_i \leq A_i \leq R_i$ の範囲の超直方体 S を、 n 個のノードで同じサイズの超直方体に分割した場合に、分割された各超直方体の処理負荷の最大値と最小値 (0 より大きな値とする) の比を p とすると、処理負荷が高い超直方体に含まれるノード数が平均の p 倍になっていると期待される。超直方体 S を n 個のノードで同じサイズの超直方体に分割した場合、1 つの属性軸上でのノードの平均数は $O(n^{1/d})$ であり、処理負荷のばらつきとノード数 n に相関は無いと考えても妥当と思われるので、データネットワーク内での検索要求の転送ホップ数 $= p^{1/d} \times d \times O(n^{1/d}) = O(n^{1/d})$ となる。

5. 今後の課題

提案方式の有効性を検証するために、OverlayWeaver[7]等のツールを用いて、シミュレーションによるメッセージ数や負荷分散の評価を行うことが必要である。

6. 謝辞

日頃より御指導頂いております通信部門 伊藤主

管技師長、松山部長に感謝致します。

参考文献

- [1] A. Bharambe, M. Agrawal, and S. Seshan. Mercury: Supporting scalable multi-attribute range queries. In Proceedings of SIGCOMM, 2004
- [2] X. Li, Y. J. Kim, R. Govindan, H. Wei, Multi-dimensional Range Queries in Sensor Networks, In Proceedings of the ACM Sensys 2003 (Nov. 2003)
- [3] I. Stoica, R. Morris, D. Karger, F. Kaashoek and H. Balakrishnan, Chord: A scalable peer-to-peer lookup service for internet applications, In Proceedings of the SIGCOMM'01 Symposium on Communications Architectures and Protocols (2001)
- [4] S. Ratnasamy, P. Francis, M. Handley, R. Karp and S. Shenker, A Scalable Content-Addressable Network, In Proceedings of ACM SIGCOMM, Aug. 2001
- [5] D. Kostic, A. Rodriguez, J. Albrecht, A. Bhirud, and A. M. Vahdat. Using random subsets to build scalable network services. In Proceedings of USITS, Seattle, WA, March 2003
- [6] N. J. A. Harvey, M. B. Jones, S. Saroiu, M. Theimer, and A. Wolman. SkipNet: A Scalable Overlay Network with Practical Locality Properties. In Fourth USENIX Symposium on Internet Technologies and Systems (USITS), 2003
- [7] <http://overlayweaver.sourceforge.net/index-j.html>