

センサデータ分散管理システムにおけるレプリケーションフレームワークの実現

三木佑真 横田 裕介 大久保 英嗣

立命館大学情報理工学部

1 はじめに

無線センサネットワークを用いた環境観測システムなどにおいては、センシングの対象エリアが拡大するほど、マルチホップ通信による電力コストの増大が大きな問題となる。この問題を解決するために、センサネットワークを適切なサイズに分割するという方法が考えられる。これをネットワーク環境がない屋外環境で実現するために、我々は P2P データポットシステムと呼ばれる無線通信機能を持つ小型計算機によって構成される、センサデータを分散管理システムを開発してきた。これは無線アドホック通信でネットワークを形成し、センサデータを分散管理するシステムである。P2P データポットは遠方の P2P データポットのセンサデータを参照する際にホップ数の増加により応答性が低下するという問題がある。これを解決するために、我々はプロアクティブコンピューティングの考え方を導入した。本稿ではこれを実現するために必要となるレプリケーションフレームワークについて述べる。

2 センサデータ分散管理システム: P2P データポット

2.1 P2P データポットの概要

P2P データポット [?] とは、無線通信機能を持つ小型計算機によって構成され、自身に直接接続されたセンサネットワークから取得したデータをストレージに蓄積する。無線アドホック通信を行い、複数の P2P データポットが協調動作することにより、複数のセンサネットワークによって収集されたセンサデータを分散管理することを可能にする。

2.2 P2P データポットの課題と解決策

P2P データポットは無線アドホック通信を用いてネットワークを構成するため、遠方の P2P データポットと通信する際にホップ数の増加によって応答性が低下するという問題がある。環境観測システム等のアプリケーションでは状態の変化をリアルタイムに観測し、迅速に予測や警告を行いたいという要求がある。そのため可能な限り短時間で環境データを提供する必要がある。この問題に対処するためには、例えばシステム内の全てのデータを各 P2P データポットがあらかじめ所持しておくことに

より、要求後のデータ転送を避ける方法が考えられるが、センサデータは更新頻度が高いため、膨大な数のデータ同期が必要となり、無線ネットワークが破綻するおそれがある。そこで我々はプロアクティブコンピューティングの考え方を導入することにより、この問題の解決を試みることにした。具体的には数回の人間の操作から次の操作を予測して自律的に動作させるという手法である。ユーザからのデータ問い合わせの履歴から次の問い合わせを予測し、必要となりそうなデータを必要となりそうな場所の付近まで事前に転送しておく [?]. このデータが使用されることにより、データ転送の時間が削減され応答性の向上が期待できる。以後、この事前に転送しておくデータのことをプロアクティブデータと呼ぶ。

2.3 プロアクティブデータ管理機構

プロアクティブデータ管理機構はプロアクティブデータの作成を行い、問い合わせが発行されたときにプロアクティブデータが使用できるかを調べ、使用できる場合は問い合わせをその場所まで転送する。これを実現するためにプロアクティブデータ管理機構はプロアクティブデータの作成、作成されたプロアクティブデータの検索、不要なプロアクティブデータを削除という 3 つの機能が become 必要になる。プロアクティブデータの作成には RDBMS のレプリケーションという機能を利用し、他の P2P データポットが持つデータベースの複製を作成し、常に内容が自動的に同期されるようにする。これらの機能をレプリケーションフレームワークとして提供する。

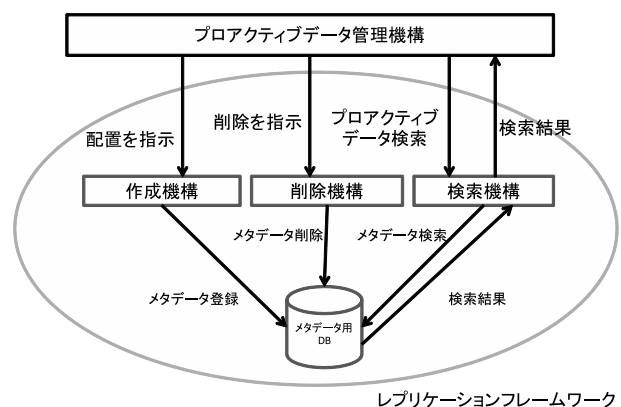


図1 レプリケーションフレームワークの構成

2.4 関連研究

P2P ネットワーク上でデータのレプリケーションを行う研究には、ファイルなどのサイズの大きな単位で行わ

れるものがある[?]。これらとは異なり、本機構はリレーショナルデータベースのテーブル単位でのデータレプリケーションを行うものである。

3 レプリケーションフレームワークの設計

レプリケーションフレームワークの構成を図1に示す。プロアクティブデータ管理機構は2.3節で述べた3つの機能と、プロアクティブデータの管理に使用するメタデータのデータベースから構成される。

3.1 メタデータ

プロアクティブデータの検索や削除などに使用するために、作成されたプロアクティブデータに「情報元」「作成先」「テーブル名」「作成時刻」「最終更新時刻」「最終読込時間」の6つのメタデータを付与する。またこのうち、「情報元」「作成先」「テーブル名」の3つについては、全てのP2Pデータポットで共有する。これにより、プロアクティブデータの検索時に他のP2Pデータポットと通信する必要がなくなる。残りの3つに関してはプロアクティブデータを所持するP2Pデータポットのみが所持する。

3.2 レプリケーション作成機構

プロアクティブデータ管理機構から指定されたP2Pデータポットにレプリケーションを作成する。作成時にメタデータを登録する。

3.3 レプリケーション検索機構

ユーザからのデータ参照クエリを解釈し、どのP2Pデータポットのデータを利用しようとしているかを特定し、メタデータを用いて使用可能なプロアクティブデータを検索する。またそれが複数存在する場合、位置情報やホップ数の情報を用いて、より近い場所に存在するものを使用する。使用可能なプロアクティブデータが存在しない場合はオリジナルのデータを使用する。

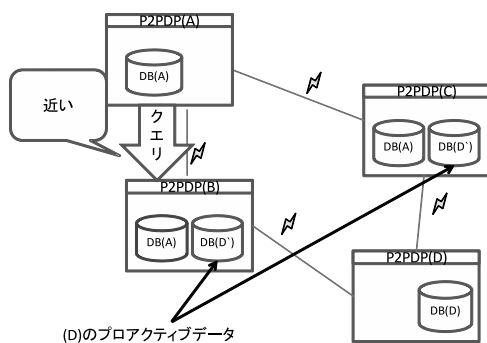


図2 検索機構の動作例

図2に具体例を示す。ここではP2Pデータポット(A)～(D)の4台が稼働しており、(D)のプロアクティブデータが(B)および(C)に配置されているものとする。この時(A)から(D)の持つセンサデータを要求するクエリが発行された場合、(A)はメタデータを利用し、(D)が持つセンサデータに対応するプロアクティブデータを検索する。その結果、使用可能なデータは(B)、(C)、(D)にあるこ

とがわかる。その3つの候補の中で最も近いデータポットにクエリを送信する。

3.4 レプリケーション削除機構

プロアクティブデータ管理機構が、プロアクティブデータを不要だと判断した場合に呼び出され、その削除を行う。

4 実装

4.1 メタデータ

P2PデータポットはP2PネットワークのフレームワークとしてJXTAを利用している。P2PデータポットはJXTAで他のP2Pデータポットを発見するためにアドバタイズメント機能を利用している。このアドバタイズメントで扱う情報として「情報元」「作成先」「テーブル名」の3つを追加し、P2Pデータポット間での情報共有を実現する。

4.2 レプリケーション作成機構の実現

P2Pデータポットはセンサデータを格納するためのデータベースとしてPostgreSQLを使用している。PostgreSQL上でレプリケーション機能を実現するための手段は複数存在するが、今回はSlony-I[?]を使用した。Slony-Iを選択した理由はテーブル単位でのレプリケーションが可能なことや、データベースを停止させることなく、稼働中のままの状態でのレプリケーションを行うことができることなどが、通信帯域やストレージ容量などが限られるP2Pデータポットの想定環境に適しているためである。

5 おわりに

本稿では、センサデータ分散管理システムであるP2Pデータポットへのレプリケーションフレームワークの実現について述べた。本機構を用いることで、P2Pデータポットの応答性を向上させることができる。今後は、プロアクティブデータの配置、検索、削除機能の動作確認、およびプロアクティブデータ管理機構を用いない場合のP2Pデータポットと比較した応答性の向上に関する評価実験を行う予定である。

参考文献

- [1] 藤崎友樹, 鈴木和久, 横田裕介, 大久保英嗣, “無線アドホック通信を利用したセンサネットワーク向け協調ストレージシステム,” 情報処理学会研究報告, 2008-MBL-44/2008-UBI-17, pp. 149-156, 2008年3月。
- [2] 井邊研吾, 横田裕介, 大久保英嗣, “センサデータ分散管理システムにおけるプロアクティブデータ配置・検索機構の性能評価,” 信学技報, vol. 111, no. 386, USN2011-73, pp. 95-100, 2012年1月。
- [3] 影山潤, 渋谷進, “P2Pネットワークにおけるファイル消失を防ぐ複製配置手法,” 第1回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2009), C9-3, 2009年3月。
- [4] Slony-I: <http://slony.info/>.