

論文 / 著書情報
Article / Book Information

題目(和文)	大規模データ処理基盤へのエージェント適用に関する研究
Title(English)	
著者(和文)	村田悠也
Author(English)	Yuya Murata
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10378号, 授与年月日:2016年12月31日, 学位の種別:課程博士, 審査員:寺野 隆雄,新田 克己,出口 弘,三宅 美博,小野 功
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第10378号, Conferred date:2016/12/31, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

専攻： Department of	知能システム科学	専攻	申請学位 (専攻分野)： 博士 Academic Degree Requested Doctor of	(工学)
学生氏名： Student's Name	村田悠也		指導教員 (主)： Academic Advisor(main)	寺野隆雄
			指導教員 (副)： Academic Advisor(sub)	

要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters)

システムから発生するデータや蓄積されるデータは年々増加し、量や発生速度からビッグデータと呼ばれ、処理基盤や利用方法が研究されている。特に、金融取引やモバイルアプリケーション、IOT、MVNO などの通信サービスで用いられるシステムは短時間に継続的かつ大量にデータが発生する。このようなシステムは、ユーザーやデバイスの増加、サービスの拡張等により短時間に発生するデータが爆発的に増加する。例えば、位置情報を利用したアプリケーションを 1 万人のユーザーが利用しているとき、毎秒 1 万件のデータがアプリケーション実行中に発生する。さらに、位置情報を他のユーザーと共有する場合、データは 1 万×1 万の秒間 1 億件まで増加する。このような短期間に大量のデータが継続的に発生するようなシステムでは、データの蓄積による処理が困難となる。そこで開発されたのが、リアルタイムデータ集計処理(Realtime Data Aggregation, RDA)と呼ばれる処理手法である。RDA は、データをリアルタイムに繰り返し集計することでデータを蓄積せずに集計処理する。しかしながら、その特徴から処理要件が非常に厳しく、その実現は難しい。一方、高いデータアクセス性能とスケーラビリティ性能を有したエージェントベースのインメモリデータストアが存在する。このインメモリデータストアは、エージェントの特徴である非同期性や反応性を用いることによりリアルタイムかつスケーラブルな処理を実現する。一方で、その設計や運用方法はシステム開発経験者の熟練の技能となっており、RDA への適用方法について十分に議論されていない。そこで本研究の目的は、秒間数万～数百万件のストリームデータをリアルタイムに集計処理するデータ処理基盤に対し、高いデータ処理性能を持つエージェント技術を適用するとともに、その実現方法と処理性能課題を解決するための効果的なエージェント設計方法論と運用技術の確立をする。

本論文は 5 章から構成され、その概要は次のとおりである。第 1 章は「序論」であり、研究の背景や目的、論文構成について述べる。第 2 章の「関連する研究領域の手法」では、近年、ビッグデータ処理に用いられている MapReduce や RDDs といった並列分散処理のプログラミングモデルについて述べ、本研究で用いるエージェント技術について解説する。第 3 章「動的エージェント負荷分散機構」では、はじめに単純な RDA システムの例としてユーザーアプリケーションのセンサーログを解析するセンサーログマイニングシステムを紹介し、エージェントの適用を行う。単純な RDA システムは、設計が容易でエージェント間の通信が発生しないことから、エージェント配置時の運用上の問題を明らかにすることができる。これにより、エージェントシステムでは、アプリケーション利用者の年齢や地域による偏りがエージェントのメッセージの集中を引き起こし、性能が大幅に劣化することを確認した。また、その対策としてエージェントがリアルタイムに変化するアプリケーション状況に合わせて数の調節や負荷分散を行う動的エージェント負荷分散機構を開発し、単一環境下でデータの発生状況を変えて評価した。第 4 章「複雑な RDA へのエージェント適用」では、複雑な集計処理であるランキングシステムへエージェントを適用する。ランキングシステムは、データの集計がユーザー個別に集計後、全体で集計する 2 段階の集計処理となっている。そのためエージェントシステムにより実現した場合、エージェント間で通信が発生する。また、センサーログマイニングシステムと異なり、ランキング時に複数のエージェントにまたがってデータを参照する必要がある。本研究では、複雑な RDA システムへのエージェント適用にあたり、エージェント設計とエージェント分散配置の 2 段階設計法を考案しシステム要件からのエージェント導出法を開発した。また、エージェント分散配置では、通信状況に応じて 2 つの配置戦略、エージェントタイプ配置とエージェントクロニグ配置を提案し、モバイルアプリケーションにおけるリアルタイムランキングを例に、その比較評価を行った。第 5 章「結論」では、本研究の貢献と課題について述べる。

本研究の貢献は、エージェントの 2 段階設計法と動的エージェント負荷分散機構の開発により、熟練の技能によらず高性能な RDA システムをエージェントにより実現することが可能となったことにある。今後の課題としてサーバーでの通信の偏りに対してエージェント側での解決方法について開発する。これにより、サーバーを接続するだけでエージェントシステムの拡張が可能となる。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。
Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).

(博士課程)
Doctoral Program

論文要旨

THESIS SUMMARY

専攻： 知能システム科学 専攻
Department of
学生氏名： 村田悠也
Student's Name

申請学位 (専攻分野)： 博士 (工学)
Academic Degree Requested Doctor of
指導教員 (主)： 寺野隆雄
Academic Advisor(main)
指導教員 (副)：
Academic Advisor(sub)

要旨 (英文 300 語程度)
Thesis Summary (approx.300 English Words)

Recently, large scale data problem called Bigdata that the development of high-performance architecture and usage has been attracting attention. Real Time Data Aggregation (RDA) is a technology for aggregating the massive amounts of data generated in a short period of time. However, RDA is the challenge the performance requirements. On the other hand, in-memory data store by the agent programming model is a parallel distributed processing technology to completely fill the requirement of the RDA. However, it is not enough discussion about the agent application to the RDA. In this study, to implement of the RDA with agent techniques, it focused on the two problems follows: a) design of the agent and b) management of the agent system. The thesis consists of five chapters as follows. The first chapter is the introduction which describes the background, purpose and result of this research. In the second chapter, we explain related works which the programming model in a parallel distributed processing and agent techniques. In the 3th chapter, we propose a simple RDA system with a sensor-log mining system. Simply RDA system is an easily implemented by the agent and able to monitor an agent load. Therefore, It clarifies the problem of the management of an agent system. We develop a dynamic agent load balancer mechanism that detects the agent load and dynamic load balancing. In the 4th chapter, to apply the agent to the complexity RDA system. Complexity RDA is difficult because the communication occurs between the agent. We propose method of the Agent Design and the Deployment Distributed Agent. The 5th chapter is the conclusion of this research. The main contributions of this research are to develop the a) method of the 2 stage design with the "Design of the Agent", and "Design of Distribution Deployment of Agents", b) Dynamically Agent Load Balancer.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。
Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。
Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).