HOKKAIDO UNIVERSITY

Title	Multi-Objective Resource Optimization of Big Data Applications in Cloud Computing Infrastructure [an abstract of dissertation and a summary of dissertation review]
Author(s)	Phyo Thandar Thant
Citation	北海道大学. 博士(情報科学) 甲第13079号
Issue Date	2018-03-22
Doc URL	http://hdl.handle.net/2115/70249
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/4.0/
Type	theses (doctoral - abstract and summary of review)
Additional Information	There are other files related to this item in HUSCAP. Check the above URL.
File Information	Phyo_Thandar_Thant_review.pdf (審査の要旨)

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学位論文審査の要旨

博士の専攻分野の名称　博士 (情報科学)　氏名　Phyo Thandar Thant

審査担当者　主　査　教　授　棟朝 雅晴
副　査　教　授　髙井 昌彰
副　査　教　授　水田 正弘
副　査　准教授　杉木 章義

学位論文題名

Multi-Objective Resource Optimization of Big Data Applications in Cloud Computing Infrastructure
（クラウドコンピューティング基盤におけるビッグデータアプリケーションの多目的資源割当最適化）

科学技術分野におけるビッグデータ処理が近年その重要性を増しており, クラウドコンピューティングの急速な発展にともない, 様々なクラウドサービスや計算資源を組み合わせて, ゲノム解析ワークフローなどのビッグデータ処理を実現することが可能となりつつある。しかしながら, 多岐にわたるサービスや資源を組み合わせて, 高性能なシステムを低コストで, かつ信頼性などの制約条件を満たしつつ構築することは困難な課題となっており, その最適化に向けた研究開発がなされてきた。

本論文では, 科学技術分野におけるビッグデータ処理において, 最も一般的に用いられるものとして, Hadoop に代表されるビッグデータ処理フレームワークと, ゲノム解析ワークフローなどのサイエンティフィックワークフローを例にとり, それらの処理をプライベートクラウドやパブリッククラウドで提供される仮想マシンを用いて効率的に実現するための, 多目的資源割当最適化について議論している。ビッグデータ処理には多くの計算資源および計算時間を要することから, 研究者の観点からは, できるだけ短い時間で処理を終えること, その処理に必要となるクラウド利用料金のコストを最小化すること, さらには実行時の信頼性についても確保する必要があるなど, 複数の評価尺度を同時に考慮した多目的最適化が求められる。また, 特にパブリッククラウドでは数多くのインスタンスタイプが提供されるなど, サービスの選択肢が多岐にわたり, また, Hadoop などのフレームワークやワークフローの構築方法などに関するパラメータも数多く存在することから, 最適なパラメータを手動で選択するのは事実上不可能である。そこで, 本論文においては, 多目的進化計算アルゴリズムを用いた統一的な枠組みのもとで, クラウドコンピューティング基盤におけるビッグデータ処理に関わる多目的資源割当最適化問題を解決する。

第 1 章では, ビッグデータ処理における多目的最適化の必要性, および本論文において提案する解決策の概要について議論している。

第 2 章では, 関連技術および既存研究として, クラウドコンピューティング, Hadoop などビッグデータ処理フレームワーク, サイエンティフィックワークフロー, 多目的進化計算アルゴリズムについて議論している。

第 3 章では, ビッグデータ処理フレームワークとして代表的な Hadoop を例にとり, その資源割当およびパラメータ最適化を, 実行時間および実行コストの観点から多目的進化計算アルゴリズム

を適用した手法について議論している。多目的進化計算アルゴリズムの適用において, その適応度評価に Hadoop を実環境で実行する必要があり多くの計算時間を必要とすることから, 適応度評価を削減するために集団内の一部の個体について評価を行い置き換えつつ探索をすすめる改良を行った。その結果として, 実行時間と実行コストのトレードオフ関係を反映したパレート最適解集合を求めることができ, Hadoop の利用者が多岐にわたる複雑なパラメータの設定を手動で行う必要なく, 効率的なビッグデータ処理を実現することができた。

第 4 章では, ゲノム解析処理フローなどのサイエンティフィックワークフローによるビッグデータ処理に関するクラウド資源割当最適化問題に対して, 多目的進化計算アルゴリズムを適用することで, 実行時間, 実行コスト, さらには SLA (Service Level Agreement) など信頼性に関する複数の評価尺度を考慮したパレート最適解集合を求めることができた。

第 5 章, 第 6 章においては, 以上の結果に関して考察するとともに, 今後に向けた展望について議論している。結果として, 提案する多目的進化計算アルゴリズムによる多目的資源割当最適化により, 科学技術分野で代表的なビッグデータ処理アプリケーションである Hadoop, およびサイエンティフィックワークフローにおいて, 複雑かつ多岐にわたるクラウド資源割当最適化およびそのパラメータ設定最適化に関わる問題を, 統一的な枠組みにより解決し, 得られたパレート最適解集合をもとに研究者がビッグデータ処理を短時間, 低コストで, かつ高い信頼性のもとで行える基盤技術を確立したものである。

これを要するに, 本論文は先端的な多目的進化計算アルゴリズムを用いることで, ビッグデータ処理フレームワークにおけるクラウド資源割当・パラメータ最適化, および, サイエンティフィックワークフローに対応するクラウド資源割当の最適化を実現しており, 近年その重要性が高まっているビッグデータ処理アプリケーションの効率的な実現に貢献するところ大なるものがある。よって著者は, 博士 (情報科学) の学位を授与される資格あるものと認める。