

Title	Performance Modeling of Large-Scale Parallel-Distributed Processing for Cloud Environment(Abstract_要旨)
Author(s)	Hirai, Tsuguhito
Citation	Kyoto University (京都大学)
Issue Date	2018-05-23
URL	https://doi.org/10.14989/doctor.k21280
Right	Chapter 2: (c) 2014 the American Institute of Mathematical Sciences. Re-used with permission from T. Hirai, H. Masuyama, S. Kasahara and Y. Takahashi, “ Performance Analysis of Large-Scale Parallel-Distributed Processing with Backup Tasks for Cloud Computing, ” Journal of Industrial and Management Optimization, vol. 10, no. 1, pp. 113–129, 2014 (DOI: 10.3934/jimo.2014.10.113). Chapter 3 and Appendix A: (c) 2016 the Operations Research Society of Japan. Re-used with permission from T. Hirai, H. Masuyama, S. Kasahara and Y. Takahashi, “ Performance Analysis of Task Replication in Large-Scale Parallel-Distributed Processing: An Extreme Value Theory Approach, ” Journal of the Operations Research Society of Japan, vol. 59, no. 2, pp. 174–194, 2016 (DOI: 10.15807/jorsj.59.174). Chapter 4: (c) 2018 the American Institute of Mathematical Sciences. Re-used with permission from T. Hirai, H. Masuyama, S. Kasahara and Y. Takahashi, “ Performance Optimization of Parallel-Distributed Processing with Checkpointing for Cloud Environment, ” Published online in Journal of Industrial and Management Optimization (DOI: 10.3934/jimo.2018014).
Type	Thesis or Dissertation
Textversion	ETD

(続紙 1)

京都大学	博士 (情報学)	氏名	平井 嗣人
論文題目	Performance Modeling of Large-Scale Parallel-Distributed Processing for Cloud Environment (クラウド環境における大規模並列分散処理の性能モデル)		
(論文内容の要旨)			
<p>クラウド・コンピューティングにおける並列分散処理では、巨大なタスクを多数のサブタスクに分割し、それぞれをワーカと呼ばれるマシン群に独立に処理させる。これにより、巨大なデータの処理を伴うタスクなどを比較的短時間で処理することができる。しかしながら、並列分散処理においては、「落伍者の問題」、すなわち、一部のワーカが故障あるいはその他の原因によりサブタスクの処理に多大な時間を要し、並列分散処理がうまく機能しない状況が発生する場合がある。落伍者の問題は、サブタスクの数が増加するにつれてより深刻なものとなる。</p> <p>落伍者の問題を低減するための手法として、タスク複製法とチェックポイント法の2つの方法が知られている。タスク複製法では、各サブタスクは複製され、オリジナルのサブタスクとその複製が、それぞれ別々のワーカに割り当てられる。サブタスクの処理は、オリジナルのサブタスクとその複製のいずれかが終了した時点で完了する。このため、タスク複製は、処理能力が低下したワーカによる影響を軽減することができる。一方、チェックポイント法では、サブタスクを処理する間、処理の進捗をチェックポイントとして定期的に保存する。ワーカ故障が発生した場合、故障したワーカに処理されていたサブタスクは、別のワーカにより最新のチェックポイントから再開される。こうして、チェックポイント法は、処理に失敗したサブタスクを最初から処理し直すことを防止する。</p> <p>先行研究では、落伍者の問題が存在する事実は指摘されているが、問題の検証を行うためのモデルについては十分な検討が行われていない。また、タスク複製法やチェックポイント法の効果や設計に関して、実機実験や確率的シミュレーションに基づく検討が行われているものの、システム環境の想定が限られており、十分な知見が得られているとはいえない。本論文では、落伍者の問題が並列分散処理の性能に及ぼす影響、ならびに、タスク複製法とチェックポイント法の効果について、解析モデルの構築とその評価を通じて検討を行っている。</p> <p>本論文は5つの章から構成され、第1章は序論である。第2章では、落伍者の問題による影響について検証するため、タスク・スケジューリング・サーバを、多数のワーカで構成される一つのサーバと見立てた待ち行列としてモデル化し、最大スループットと平均応答時間を解析的に導出している。また、数値実験によって、ワーカ処理時間の変動係数が大きくなるにつれ、落伍者の問題による性能低減が増大することが確認されている。</p> <p>第3章では、タスク複製法の効果を検証するため、極値統計理論を用いて、タスク処理時間の平均と分散の近似値をタスク複製数の関数として導出している。その導出結果に対する考察と数値実験を通じて、タスク複製法によるタスク処理時間の短縮率は、ワーカの処理時間分布の裾に大きく依存することが示されている。</p> <p>第4章では、タスクの処理時間に対するチェックポイント法の効果を評価するため、チェックポイント法を具備した並列分散処理の解析モデルを構築し、タスクの平均処理時間とこれを最小化するチェックポイント数に関する陽な近似式を導出している。特に、後者の近似式は、単体処理に対する先行研究の結果を並列分散処理に一般化したものである。数値実験では、導出された近似式がタスクの平均処理時間の評価、および最小化に有用であることが示されている。</p> <p>最後に、第5章では、論文全体のまとめとともに今後の研究の方向性を述べてい</p>			

る。

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し
審査結果の要旨は日本語500～2,000字程度で作成すること。

(続紙 2)

(論文審査の結果の要旨)

クラウド環境下の大規模並列分散処理では、処理性能の遅い一部のワーカが全体の処理時間を大きく遅延させる「落伍者の問題」が発生する。本論文は、落伍者の問題に起因するシステムの性能劣化を軽減するための代表的な手法であるタスク複製法とチェックポイント法について、待ち行列理論および極値統計理論を用いた性能評価モデルを構築し、その解析結果に基づいて、大規模並列分散処理システムの設計に有益な知見を得ている。主な結果は以下の通りである。

1. サーバでの輻輳遅延を考慮した並列分散処理の性能評価を行うため、待ち行列モデルを構築している。さらに、構築した待ち行列モデルの解析を通して、サブタスクの処理を行うワーカの処理時間分布の変動係数が、落伍者の問題に起因するシステムの性能劣化の程度に対して支配的であること、加えて、タスク複製法による性能向上の程度が、ワーカの処理時間分布の形状に依存することを明らかにしている。
2. ワーカでの処理時間のみに着目した並列分散処理の解析モデルを新たに構築し、タスク複製法におけるタスク複製数が、タスク処理時間に与える影響について検討している。本解析モデルに対して、極値統計理論に基づく漸近解析を行い、タスク複製法の効率、および、タスク複製数に対する平均タスク処理時間の感度が、ワーカの処理時間分布の裾減衰速度に大きく依存することを明らかにしている。
3. チェックポイント法による落伍者の問題の低減効果を明らかにするため、ワーカでの処理時間の解析モデルを構築している。さらに、チェックポイント法を特徴づけるパラメータ間の相対的な大小関係を利用して、処理時間分布の近似を行い、平均タスク処理時間および最適なチェックポイント回数に関する簡便な近似式をそれぞれ導出している。特に、後者の近似式は、並列分散処理に対応していない既存の近似公式の一般化となっている。また、数値実験によって、導出した近似式に基づいてチェックポイント回数を適切に設定することでシステムの性能を大きく改善できることを明らかにしている。

クラウド環境における大規模並列分散処理で処理性能の低下の要因となる落伍者の問題に注目し、大規模並列分散処理のタスク処理時間に対して待ち行列の観点から数理モデル化を行うとともに、タスク処理時間の解析的な評価を通じて、ワーカ処理時間分布の性質やワーカ数などの要因が並列分散処理によるタスク処理効率にどう影響するかを具体的に明らかにした本論文の成果は、大規模並列分散処理の効率化に際して有用な知見を与えるものである。また、サブタスク数が大きい状況でのタスク処理時間の評価に極値統計理論の適用を提案していることについては、申請者の着想の独自性も認められる。以上により、本論文は大規模並列分散処理における処理性能の評価ならびに効率化に関して新規かつ有用な研究成果を取りまとめた論文として情報学分野における学術的な貢献度が高く、したがって博士(情報学)の学位に値するものと認められる。なお、平成30年4月16日に、本論文の内容ならびに関連した事項についての試問を行った結果、合格と認めた。

注)論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。

更に、試問の結果の要旨（例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」）を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。

要旨公開可能日： 年 月 日以降