

|           |  |         |       |
|-----------|--|---------|-------|
| 氏 名       | 李 燮鳴   |         |       |
| 学 位 の 種 類 | 博 士 (工 学)  |         |       |
| 学 位 記 番 号 | 博 甲 第 8077 号   |         |       |
| 学位授与年月日   | 平成 29 年 3 月 24 日   |         |       |
| 学位授与の要件   | 学位規則第 4 条第 1 項該当   |         |       |
| 審 査 研 究 科 | システム情報工学研究科  |         |       |
| 学位論文題目    | A Study on Data-Aware Scheduling for Post-Peta Scale Systems<br>(ポストペタスケールシステムにおけるデータ配置を考慮した<br>タスクスケジューリング) |         |       |
| 主 査       | 筑波大学 教授  | 博士 (理学) | 建部 修見 |
| 副 査       | 筑波大学 教授  | 工学博士    | 朴 泰祐  |
| 副 査       | 筑波大学 教授  | 博士 (工学) | 追川 修一 |
| 副 査       | 国立情報学研究所 教授  | 博士 (工学) | 合田 憲人 |
| 副 査       | 筑波大学 准教授   | 博士 (工学) | 川島 英之 |

## 論 文 の 要 旨

本論文では、データの局所性を利用するバッチキューイングシステムのスケジューリングの研究を行っている。データが各計算ノードに配置されている場合は、データの局所性を利用してジョブを配置することにより I/O 処理時間を短縮することができ、ジョブの実行時間を短縮することができる。これまでバッチキューイングシステムのスケジューラは CPU の利用率を上げることが大きな目的であったが、ジョブの実行時間の短縮のためにデータの局所性を利用することが求められる。データの局所性を利用するスケジューラとしては、既存研究として MapReduce のスケジューラがある。MapReduce では、Map タスクは単一データを入力とするため、データの格納されている計算ノードを自然に割り当てることができる。一方、バッチキューイングシステムにおいては、汎用的なジョブが実行されるため、入力データは複数ファイルになることがある。そのため、本研究ではファイル局所性という指標を導入し、ファイル局所性と CPU 負荷をパラメータ  $\beta$  で組み合わせたスコアを元にスケジューリングを行う Data-Aware Dispatch (DAD) を提案している。また、ジョブがいかにデータインテンシブであるかを示す指標として Remote Degradation Rate (RDR) を提案し、RDR と遠隔のファイルサイズの比を元にしてスケジューリングを行う Improving Data-Aware Dispatch (IDAD) を提案している。また、データの局所性を有効に使うため、MapReduce のスケジューラで提案されている遅延スケジューリング (Delay Scheduling) を DAD, IDAD に応用している。MapReduce では、入力は単一データであるため、データが存在するかないかの状態しかないが、DAD, IDAD ではその中間状態があるため、スコアをベースに局所性があるかどうかを定める必要がある。DNA の塩基配列やたんぱく質のアミノ酸配列の相同性検索のベンチマークである BLAST ベンチマークにおいて、30%以上の性能向上を達成した。

## 審 査 の 要 旨

### 【批評】

データインテンシブコンピューティングにおいては、計算ノードのローカルストレージの活用が有望である。そのため、バッチキューイングシステムにおけるスケジューリングについてもデータ局所性を活用するものが求められている。これまでデータ局所性を活用するものとして、MapReduce のスケジューラがあったが、MapReduce の実行では Map タスクは単一データを入力とするものであり、一般的なバッチキューイングシステムにそのまま適用することはできない。本研究では、バッチキューイングシステムのスケジューリングに適用するため、ファイル局所性の新たな指標の提案を行っている。MapReduce では、単一データを入力とするため、局所性があるか、ないかのどちらかであるが、バッチキューイングシステムにおけるジョブにおいては、複数の入力をとることもあり、ファイル局所性を表す新たな定義が必要となる。また、ジョブによっては、CPU インテンシブなジョブと I/O インテンシブなジョブがある。これらを区別するために、本研究では全てのデータについて局所性がある場合と、ない場合の実行差を元に Remote Degradation Rate (RDR)を提案している。本研究では、ファイル局所性を表す指標と CPU 負荷をパラメータ  $\beta$  で結合したスコアによる方法(DAD)と、RDR と局所性のないファイルの割合による方法(IDAD)の提案を行っている。また、局所性を有効に活用するため MapReduce のスケジューラで提案されている遅延スケジューリング(Delay Scheduling)のバッチキューイングシステムに対する適用も提案している。オープンソースのバッチキューイングシステムである Torque に対し実装を行い、評価については、ファイルを参照するマイクロベンチマークと、DNA の塩基配列やたんぱく質のアミノ酸配列の相同性検索ベンチマークである BLAST ベンチマークにより行っている。BLAST ベンチマークは、CPU インテンシブなジョブと I/O インテンシブなジョブがミックスしたベンチマークである。このようなベンチマークの場合、DAD では  $\beta$  の決定と遅延スケジューリングの遅延の決定が難しい。評価においては、複数の  $\beta$  と遅延で評価を行い、最もいいものを選ぶ必要がある。IDAD では、 $\beta$  のようなパラメータの調整は必要ない代わりに、RDR の計算のために一度ジョブを実行する必要がある。評価環境における BLAST ベンチマークの評価結果では、DAD、IDAD とともに 30%を超える性能向上を達成している。バッチキューイングシステムにおけるファイル局所性を有効に利用するためにスケジューリングの提案を行い、実際のバッチキューイングシステムに実装し、アプリケーションベンチマークで評価し、有効性を示したことは評価に値する。

### 【最終試験の結果】

平成 29 年 2 月 9 日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

### 【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。