

別紙様式 2

図書館情報メディア系リサーチグループ研究成果報告書

提出日 平成26年 8月18日

リサーチグループの名称	
ネットワーク連携型 Hadoop スケジューリング(Network-Aware Hadoop Scheduling)	
リサーチグループ設置期間	
平成25年7月 ～ 平成26年6月	
リサーチグループ構成員	
所 属 ・ 職 名	氏 名
図書館情報メディア系・教授	川原崎 雅敏
図書館情報メディア系・教授	杉本 重雄
図書館情報メディア系・准教授	阪口 哲男
研究目的	
<p>本研究は、ビッグデータを多数の計算機で並列処理する Hadoop タスクスケジューリングの性能向上に関するものである。現行の Hadoop は、計算機の負荷状態だけを見てタスク割り当てを行い、計算機間のデータ転送に伴う網の過負荷を意識しないため、網輻輳が発生するとジョブ処理性能が低下することが報告されているが、詳細は明確になっていない。そこで、網状態を意識したスケジューリング方式を開発し、実機およびクラウドで構成する実験網で方式評価を行う。</p>	
研究成果	
<p>7 台の計算機からなる実機の実験網と、Amazon EC2 を用いた 20 台 (及び 50 台) の仮想マシンからなるクラウドの実験網を構築して Hadoop を実装し、そこに複数のジョブを投入して、データ転送によってジョブ処理性能が低下する現象を確認した。その実験データを分析することにより、性能低下のメカニズムを解明すると共に、それを防止する制御方式を考案し、実験網に実装して評価を行った。さらに、より大規模網での評価を可能とするために、NS3 (Network Simulator-3) を用いた Hadoop 動作解析シミュレータを開発した。これらを論文にまとめ、国内研究会および国際会議に投稿し発表を行った。</p>	

代表的な研究発表・特許等の成果一覧、特記事項等

1. H. Watanabe, M. Kawarasaki: "Impact of Data Transfer to Hadoop Job Performance: - Architectural Analysis and Experiment -", ACMSE 2014 (52nd Annual ACM Southeast Conference), 6 pages, 2014. 3, Kennesaw, USA
2. 渡邊飛雄馬、川原崎雅敏 : 「データ転送量を考慮した Hadoop タスクスケジューラ改善方式の提案と評価」、信学技報 Vol.113, No.473, pp.175-180, 2014. 3
3. 田中大地、渡邊飛雄馬、川原崎雅敏 : 「NS-3 を用いた Hadoop 動作解析シミュレータの開発」、信学技報 Vol.113, No.473, pp.181-186, 2014. 3