

氏名(本籍)	お ^お じま ^{じま} よし ^{よし} のり ^{のり} 小島好紀(茨城県)		
学位の種類	博 士(工 学)		
学位記番号	博 甲 第 4278 号		
学位授与年月日	平成 19 年 3 月 23 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	OpenMP 向けソフトウェア分散共有メモリシステムの最適化に関する研究		
主 査	筑波大学教授	理学博士	佐 藤 三 久
副 査	筑波大学教授	工学博士	朴 泰 祐
副 査	筑波大学教授	学術博士	和 田 耕 一
副 査	筑波大学教授	博士(工学)	山 口 喜 教
副 査	東京大学教授	工学博士	石 川 裕

論 文 の 内 容 の 要 旨

近年、マイクロプロセッサやネットワーク技術の進歩により、ワークステーションや PC をネットワーク結合したクラスタシステムが並列プラットフォームとして主流となっている。クラスタシステムは分散メモリ型システムであるため、一般にメッセージ通信ライブラリを用いてプログラミングが行われるが、その場合プログラミングコストが高いという問題点がある。一方、共有メモリシステムでは OpenMP で並列化することができ、並列化のコストを低く抑えられる。そこで、分散メモリシステムの上で共有メモリプログラミングを可能にするために、ソフトウェアで共有メモリを実現するソフトウェア分散共有メモリ (DSM) システムが研究、開発されている。

従来のソフトウェア DSM システムの多くは通信レイヤを TCP/IP や UDP/IP を利用して実装されており、そのプロトコルオーバーヘッドのために高速なクラスタ用ネットワークの性能を十分に活用できないことがある。また専用のプロトコルを用いて実装しているシステムもあるが、それらは可搬性の点で問題がある。そこで本論文では高可搬性と高性能を同時に実現するために、クラスタ用ネットワークで標準の通信ライブラリとして提供されている MPI を通信レイヤに用いたソフトウェア DSM システムを実装した。

ソフトウェア DSM システムでは、データの管理を各プロセッサにどのように割り当てるかがプログラムの性能に大きく影響する。本論文では、静的な割り当てでは複雑な割り当てが必要なアプリケーションに適切に対応するのが難しいという問題点を解決するために、動的な割り当て手法の 1 つである First Touch page allocation を実装し、性能評価を行った。さらに、計算とバリア同期の 2 つのフェーズそれぞれに対して投機的に通信を行う機能を実装した。計算フェーズに対してはリモートページを通信スレッドにおいてアクセス履歴を用いて Prefetch する機能を、バリア同期フェーズに対しては更新されたページデータをバリア同期ポイントに到達する前に投機的にホームノードへ送信する機能を実装した。姫野ベンチを用いた性能評価の結果、First Touch 方式では静的に最適化を行ったものより良い性能が得られた。また NPB BT に投機的通信を適用した結果、Prefetch によって 31%、投機的 diff 転送によって 3% 性能が向上した。

審 査 の 結 果 の 要 旨

広く普及している PC クラスタなどの分散メモリシステムにおいて並列プログラミングのコストを低減する方法としてソフトウェア分散共有メモリシステムがあるが、その性能を改善する技法と有効性を示した。通信スレッドによりアクセス履歴を用いて prefetch する方式では大幅な性能改善が達成されており評価できる。First Touch 方式が静的に最適化する方法よりもよい場合があるのは興味深い。残念ながら、バリア同期前にページ転送する方式については効果が得られておらず、これからの更なる研究に期待したい。また、ソフトウェア分散共有メモリの可搬性を高める実装方法として、通信レイヤに広く高性能システムで利用可能な MPI を用いるシステムを提案、実装しており、これによりソフトウェア分散共有メモリシステムの普及への貢献が期待できる。

よって、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。