

審査の結果の要旨

氏 名 遠藤 亘

本論文は、「A Decentralized Implementation of Software Distributed Shared Memory (ソフトウェア分散共有メモリの非集中型実装)」と題し、スーパーコンピュータにおけるプログラミングの生産性向上を目的としたソフトウェア分散共有メモリシステムを開発して評価した論文である。全体は7章から構成され、英文で記述されている。

第1章「Introduction (序論)」では、アプリケーションとソフトウェアシステム双方の生産性・移植性・性能などの観点から、ページベースの分散共有メモリシステムが依然として有望であると述べ、一方で従来から性能面の課題が多いことに触れている。その上で、従来型の分散共有メモリシステムでは、ノード間インターコネクトの高速化機能であるRemote Direct Memory Access (RDMA)や、近年主流となったマルチコアプロセッサへの対応が不十分で、これらに特化した新たなコヒーレンスプロトコルとその実装方式を提案すると言及している。

第2章「Background (研究背景)」では、その後の章で必要となる各技術要素について簡潔に解説している。具体的には、分散メモリ・共有メモリ・Partitioned Global Address Space (PGAS) といった各種メモリインターフェイスの違い、Message Passing Interface (MPI) のような従来型の通信インターフェイスに存在しなかったRDMAプログラミング特有の技術的制約、共有メモリシステムにとって重要なメモリコンシステンシモデル、ミューテックスやスレッドの実装方式といった広汎な技術について述べている。

第3章「MENPS: A decentralized distributed shared memory library designed for modern interconnects (MENPS: 現代的なインターコネクト用に設計された非集中型分散共有メモリライブラリ)」では、第2章で解説したRDMAプログラミング特有の制約を前提としながら、ホームマイグレーションと論理タイムスタンプを積極的に活用した新たなコヒーレンスプロトコルを提案し、それに基づく分散共有メモリライブラリMENPSについて述べている。また、コールスタックを含めた全てのメモリ領域を共有可能にすることで、OpenMPプログラムを分散共有メモリ上で動作させる際に特殊なコード変換なしに実装可能であることを示している。評価には主にベンチマークスイートであるNAS Parallel Benchmarksを用い、複数の共有メモリアプリケーションに関して軽微な変更のみで複数ノードでの性能向上を達成している他、既存手法や既存システムと比べても

性能向上したことを示している。

第4章「MECOM: a communication library using parallelized software communication offloading (MECOM: 並列化ソフトウェアオフローディングを用いた通信ライブラリ)」では、分散メモリ型かつマルチコアの環境におけるマルチスレッド性能向上を目的として、ソフトウェアオフローディングと呼ばれる手法をユーザレベルスレッドによって実装する方式を提案している。ノード間通信処理をユーザ側の計算から専用スレッドに効率よく分離することによって、マイクロベンチマークにおいて通信のオーバーヘッドとメッセージレート向上を達成したと示している。この方式はノード間通信ライブラリ MECOM として実装され、分散共有メモリライブラリの評価に組み込まれており、第3章においてMPIを使用するよりもMECOMを介してRDMAを直接使用した方が分散共有メモリの性能が向上することも示している。

第5章「ComposableThreads: a user-level threading library with compile-time parametricity (ComposableThreads: コンパイル時パラメトリシティを持つユーザレベルスレッドライブラリ)」では、追加のオーバーヘッドなしで拡張可能なユーザレベルスレッドライブラリの設計と実装について解説している。提案されているユーザレベルスレッドライブラリに関しては、分散共有メモリ上のOpenMPインターフェイスの実装に貢献している他、マイクロベンチマーク単体での性能においても、拡張可能な実装方式を採用しながら既存システムと同等以上の性能であることを示している。

第6章「Related work (関連研究)」において、ソフトウェア分散共有メモリに関する研究を中心に、ノード間通信やユーザレベルスレッド処理系についてなど広汎な関連研究を網羅的に説明している。既存の分散共有メモリ処理系については歴史的に特徴的な処理系から出発し、特に本論文にとって関連の深いRDMAベースの分散共有メモリについても提案処理系との差分を説明している。また、PGASやハードウェア共有メモリといった異なる形態のシステムについても、ソフトウェア分散共有メモリの研究と対比しながら解説している。

第7章「Conclusions (結論)」では、分散共有メモリシステムを一通り実装し評価した結論として、既存研究よりもRDMAやマルチコアプロセッサを積極的に活用し、現代的なソフトウェア共有メモリシステムのプロトタイプを実現したと総括している。その上で、本論文における今後の発展性についても言及している。

以上これを要するに、本論文は、並列分散型計算機でのアプリケーション生産性と性能の両立という重要な課題を解決するため、分散共有メモリという普遍性の高いアイデアに基づいて、メモリアーキテクチャやマルチスレッドといった様々な技術要素を考慮しながら改善を試みており、そのため、今後ますます重要性を増すと見込まれる並列分散型計算機に関するソフトウェア・ハードウェア研究の多方面に渡って影響が大きく、電子情報学の発展に寄与するところが少なくない。よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。