

ポストペタスケール高性能計算に資するシステムソフトウェア技術の  
創出

平成 22 年度採択研究代表者

H25 年度 実績報告
----------------

堀 敦史

(独)理化学研究所 計算科学研究機構  
上級研究員

メニーコア混在型並列計算機用基盤ソフトウェア

## § 1. 研究実施体制

### (1) 堀グループ

- ① 研究代表者:堀 敦史 ((独)理化学研究所計算科学研究機構、上級研究員)
- ② 研究項目
  - ・メニーコア用 OS カーネルの開発
  - ・スケーラブル並列ファイルシステム
  - ・超軽量マルチスレッド機構

### (2) 並木グループ

- ① 主たる共同研究者:並木 美太郎 (東京農工大学大学院工学研究院、教授)
- ② 研究項目
  - ・メニーコア用 OS における資源管理と仮想化方式

### (3) Dongarra グループ

- ① 主たる共同研究者:Jack Dongarra (テネシー大学、教授)
- ② 研究項目
  - ・故障レジリエンス基盤

## § 2. 研究実施の概要

本研究は、図 1 に示すような、メニーコア CPU とマルチコア CPU が混在するアーキテクチャを持つ、計算ノードが高速ネットワークにより接続されている分散メモリ型並列計算機を対象とした、システムソフトウェアの研究開発を目指している。このような混在型アーキテクチャは消費電力対性能比に優れ、将来のポストペタスケール時代の計算機アーキテクチャとして有望と考えられている。

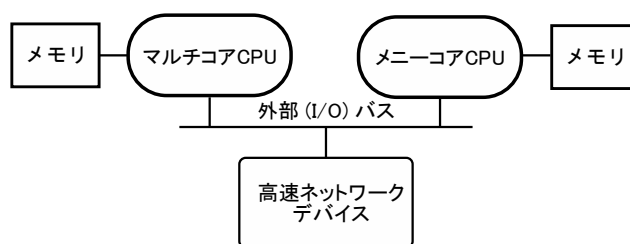


図 1 ターゲットアーキテクチャ

混在型アーキテクチャの計算ノードでは、従来型のマルチコア CPU とメニーコア CPU の密で効率的な連携が、アプリケーション実行性能の焦点となる。本研究では、マルチコア CPU が持つメモリ空間とメニーコア CPU が持つメモリ空間を統一的にひとつの仮想アドレス空間に実現し、データの受け渡しにデータのコピーを回避することで、効率的な実行モデル Multiple-PVAS (M-PVAS) を実現した(図 2)。この M-PVAS は、そもそもメニーコア CPU における、タスク間の通信を効率的に実現可能とする Partitioned Virtual Address Space (PVAS) をベースとしている。

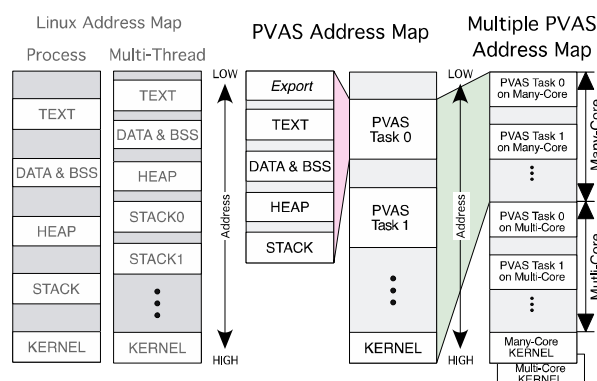


図 2 PVAS と M-PVAS のアドレスマップ

本年度は 5 年間に渡るプロジェクトの中間になる。今年度は、PVAS および M-PVAS さらにはそれらの抽象化で動作する Agent ライブラリのプロトタイプ実装がほぼ完成し、基礎的な評価をおこなった。プロジェクトの後半に向け、改良を重ね、アプリケーションレベルでの評価をおこなう予定である。

同時に、大規模なファイル I/O における問題点の明確化とその解消を目指した研究や、ポストペタにおいて予想される高頻度の故障発生に対処するための耐故障レジリエンス基盤の研究開発もおこなった。これらの研究も、プロジェクト後半にむけ、それぞれの研究を深化させる予定である。

### § 3. 成果発表等

#### (3-1) 原著論文発表

##### 論文詳細情報(国際)

1. Balazs Gerofi, Akio Shimada, Atsushi Hori, Yutaka Ishikawa, “Partially Separated Page Tables for Efficient Operating System Assisted Hierarchical Memory Management on Heterogeneous Architectures,” In CCGRID, 2013.
2. Akio Shimada, Balazs Gerofi, Atsushi Hori, Yutaka Ishikawa, “Proposing a new task model towards many-core architecture,” In Proceedings of the First International Workshop on Many-core Embedded Systems, ACM, 2013. (DOI:10.1109/PDP.2013.43)
3. Bland, W., Bouteiller, A., Herault, T., Bosilca, G., Dongarra, J.J. “Post-failure recovery of MPI communication capability: Design and rationale,” International Journal of High Performance Computing Applications August 2013 27: 244-254, doi:10.1177/109434201348823.
4. Bland, W., Bouteiller, A., Herault, T., Hursey, J., Bosilca, G., Dongarra, J.J., “An Evaluation of User-Level Failure Mitigation Support in MPI,” Computing, Springer, 2013, issn 0010-4885X, <http://dx.doi.org/10.1007/s00607-013-0331-3>.
5. Bland, W., “Toward Message Passing Failure Management,” PhD thesis, University of Tennessee, Knoxville, 2013.
6. M. Sato, G. Fukazawa, K. Yoshinaga, Y. Tsujita, A. Hori, M. Namiki, “A Hybrid Operating System for a Computing Node with Multi-Core and Many-Core Processors,” In International Journal Advanced Computer Science (IJACSci), Vol. 3, 2013.
7. Akio Shimada, Atsushi Hori, Yutaka Ishikawa, “User-Level Process towards Exascale Systems,” ROSS 2014, Munich, Germany, June 10, 2014. (submitted)