

朴 泰祐

筑波大学システム情報系・教授

ポストペタスケール時代に向けた演算加速機構・通信機構統合環境の研究開発

§1. 研究実施体制

(1) 「TCA ソフトウェア開発」グループ

- ① 研究代表者: 朴 泰祐 (筑波大学システム情報系、教授)
- ② 研究項目
 - ・ TCA 機構の FPGA ファームウェアの開発
 - ・ TCA テストボード用ホスト側ドライバの開発
 - ・ TCA 機構のユーザ API の基本設計

(2) 「TCA ハードウェア開発」グループ

- ① 主たる共同研究者: 天野 英晴 (慶應義塾大学工学部、教授)
- ② 研究項目
 - ・ TCA ハードウェアの FPGA 機構の改良に関する検討
 - ・ TCA ハードウェアの物理性能評価

(3) 「並列言語 XMP-dev/TCA 開発」グループ

- ① 主たる共同研究者: 村井 均 (理化学研究所計算科学研究機構、研究員)
- ② 研究項目
 - ・ 並列言語 XMP-dev の TCA 向け拡張の基本設計
 - ・ XMP-dev の OpenACC 対応の検討
 - ・ TCA システムソフトウェアとのユーザ API 協調設計

(4) 「TCA 向け計算科学アプリケーション開発」グループ

- ① 主たる共同研究者: 梅村 雅之 (筑波大学数理物質系、教授)
- ② 研究項目

- TCA 機構向け重力多体計算及び輻射流体計算コードの開発
- TCA 機構向け素粒子分野の格子 QCD コードの開発
- TCA 機構向け気象学分野の都市型気象計算コード (LES) 開発
- TCA 機構向け物質・生命科学分野の Car-Parrinello 分子動力学法及びメタダイナミクス法に関するコード開発

§ 2. 研究実施内容

(1) TCA システムソフトウェアの開発

- 研究のねらい

TCA 機構の実装系である PEACH2 ハードウェア上の FPGA 上で動作するファームウェアの開発、ホスト側のドライバ、ユーザ API 及び通信ライブラリ群を開発することにより、PEACH2 に基づく TCA 機構による演算加速・通信機構の融合システムの基本的実装を行い、各種アプリケーションを並列実行可能な環境を提供する。他の各グループと連携し、システム開発の中核となり、ハードウェアからアプリケーションまでを TCA コンセプト上で一貫して構築する基礎部分を構築する。

- 研究進捗状況

評価用プロトタイプ of PEACH2 ボード上で FPGA の基本設計及びその中で動作するマルチコアプロセッサ上の基本ファームウェアの設計を行った。また、ファームウェアなしの状態でも基本的な GPU 間通信を実現可能とするミニマム機能を FPGA 実装した。さらに、これらの機能を制御するためのホスト CPU 側のデバイスドライバの開発と基礎的な簡易ユーザ API を実装し、性能評価環境を整え、予備的な通信性能評価を行った。

FPGA 上の機能のみを用い、ファームウェアなしの状態での GPU・GPU 間の CPU 機能を介さないノード間直接通信性能の評価の結果、約 2.5 μ 秒の低レイテンシ通信が確認された。PEACH2 上のハードウェアレイテンシはさらに小さいが、現段階の NVIDIA 製 GPU (M2090 及び K20) では BIOS の問題、またハブとして用いている Intel SandyBridge のハードウェア特性の問題により、GPU からのデータ読み出しバンド幅が低く、レイテンシが下がらない問題を確認した。本実装により、PEACH2 ハードウェアの機能が本研究を実施するための基本仕様を満たすことが確認された [12][13]。

- 今後の予定

平成 25 年度では、別途予算により調達した PEACH2 プロトタイプボードの量産型システム及びこれを収容する中規模 GPU クラスタが利用可能となる。今年度研究成果を踏まえ、本格的なファームウェア開発と多様なユーザ要求を満たすデバイスドライバ、さらに GPU 間基本通信及び集団通信を効率的に処理する基本ライブラリを開発する。ユーザ API として CUDA 5.0 との互換性を保つレイヤを構築し、ユーザが容易にプログラミング可能な環境を提供する。

(2) TCA アーキテクチャの開発

- 研究のねらい

PEACH2-FPGA のハードウェア機構に基づき、今後の性能向上のために現在の基本機能に加えアプリケーション加速及び高度な GPU 間通信機能に資する拡張機構、オフロード機構を開発し、PEACH2 ハードウェアの改良を行う。アプリケーションの要求にハードウェアを近づけ、相互デザイン的なアプローチにより演算加速・通信機構の理想的な統合を支えるハードウェアを開発する。

- 研究進捗状況

現在の PEACH2-FPGA の設計を見直し、転送を制御する機構が持つ高速・低速 2 種類のバスの拡張モジュールとして、目的用途に応じてモジュールを準備し、転送パケットの内容・バッファの制御を行いながら処理を実装する方針を固めた。オフロード機構としては、CRC チェック機構のオフロード及びバイナリ変換機構のオフロードを実装することに決定した。CRC チェック及び生成のためのハードウェアを TCA アーキテクチャ内に組み込むことにより、ホストの負担を減らすことができる。従来、様々なハードウェアが提案されているが、ビット幅、要求訂正能力に応じて適切な処理機構を生成することができる新しい CRC アクセラレータを提案し、Xilinx 社の FPGA 上に実装した[2]。また、別途開発した独自オフロードフレームワーク”Courier”を適用し、CPU 上で実行中のバイナリから動的にデータフローを抽出し、ユーザの指示した部分を GPU 等のアクセラレータに自動的にオフロードするフレームワークの適用を検討した。データフローの抽出とアクセラレータへのデータ転送の手法を提案し、評価を行った[9]。

- 今後の予定

これらの機構を CPU からオフロードすることにより、CPU-アクセラレータ間で本質的でないデータのコピーの発生を防ぎ、オーバーヘッドをなくすることができる。また、PEACH2-FPGA 以外に GPU を搭載したヘテロジニアス構成を取る TCA での、効率的な役割分担も実現される。

(3) 並列言語 XMP-dev/TCA の開発

- 研究のねらい

TCA アーキテクチャにより低レイテンシな並列演算加速計算環境が実現されることが期待されるが、一方で TCA アーキテクチャを一般のユーザがプログラミングすることは非常に難しいと考えられる。PGAS モデル言語である XscalableMP (XMP) を GPU に適用させた XMP-dev 言語(筑波大学で開発中)を、TCA アーキテクチャ向けにさらに発展させることにより、演算加速装置と通信機構の強連結モデルに基づく strong scaling 性の高い科学技術計算アプリケーションの記述を容易にすることを目指す。

- 研究進捗状況

今年度は、主に XMP-dev/TCA の言語仕様の検討および処理系の設計を実施した。XMP-dev/TCA 言語仕様において、演算加速装置の記述仕様として OpenACC に従うことを決めた。OpenACC は今後の演算加速プログラムの中心になると予想され、これを XMP-dev/TCA に積極的に取り込む。言語処理系の基本設計として、XMP のグローバルデータのデータ管理ディスクリプタの改良、reflect, gmove, reduction 等のデータ通信指示文に対応する通信ランタイムの見直しを行い、データ分散ディスクリプタの情報に基づき、実行時に最適なグローバル通信のスケジュールを決定する方針とした。XMP-dev/TCA では、これらの通信ランタイムは、データ分散ディスクリプタに加え、データ割付けディスクリプタの情報も参照し、最適なグローバル通信および TCA 通信のスケジュールを決定する。

- 今後の予定

平成 25 年度研究では、XMP-dev/TCA 言語の仕様を確定し、プロトタイプ処理系の開発を進め、

予備評価を実施する。また、計算基礎科学分野、気象、生命科学からターゲットアプリケーションを選定し、XMP-dev/TCA による並列化を行う。さらに、XMP-dev/TCA のためのプログラミング環境の一環として、性能プロファイリングツール Scalasca を拡張することにより、XMP-dev/TCA プログラムの性能プロファイルを可能とする。

(4) TCA アーキテクチャに基づく計算科学アプリケーションの開発

- 研究のねらい

TCA アーキテクチャ構想に基づき、演算加速機構と通信機構が物理的近距离に配置され、通信レイテンシが大幅に短縮されることを想定し、これに向けた宇宙分野の重力多体計算、輻射流体計算、素粒子分野の格子 QCD 計算、気象学分野の都市型気象計算コード (LES)、物質・生命科学分野の Car-Parrinello 分子動力学法、メタダイナミクス法等の各種計算科学アプリケーションの開発を行う。

- 研究進捗状況

宇宙分野:宇宙物理学における ray-tracing 法に基づく輻射輸送シミュレーションコードを GPU を用いて大幅に高速化することに成功し、ノード並列化も達成した。また、通常は膨大な計算量の為無視される輻射性再結合による再結合光子の輻射輸送計算を GPU を用いて高速化することにも成功した[1][11]。

素粒子分野:科学的課題として、「格子 QCD による原子核の直接構成とその諸性質解明」と「有限温度・有限密度 QCD における相構造解析」を目指している。前者の計算における GPU 加速部分は大規模疎行列線形方程式の求解であり、後者では密行列の行列・行列積演算である。HA-PACS を用いてマルチノードマルチ GPU 対応のシミュレーションプログラムを開発し、特に TCA アーキテクチャによって性能向上が期待されているノード間通信の性能解析を行った。

気象学分野:建物解像 LES 気象モデルを PGI CUDA Fortran を利用して GPU 化を進め、HA-PACS 上で性能評価した。関数単体での高速化率は、4GPU 対 16CPU-core で 0.4~25.1 倍の高速化を得ており、流体計算のステンシル計算部分における袖領域通信に TCA を適用した場合の見積もりを立てた。

物質・生命科学: GPU における量子系シミュレーションコードの開発を進めた。また、GPU 分子動力学法コードの研究開発 (Platypus) を進め、HAPACS 上で、100 万原子系のタンパク質のシミュレーションを 100ns まで実行可能とした。また、GPU 分子動力学計算(NAMD)の DNA イソメラーゼに関する計算を実行した[3][4][5][6][7][8]。

- 今後の予定

今年度研究では、TCA 機構のハードウェア及びソフトウェアが完成していないことから、主としてアプリケーションの通信部分以外を GPU 化することを主目的とした。平成 25 年度以降、TCA のプロトタイプボード及びこれを用いた本格的実験環境が整備されることから、TCA のユーザ API を用いた低レイテンシ通信アルゴリズムを開発し、従来手法との差異を性能評価しつつ、アルゴリズムの改良を進める。

§3. 成果発表等

(3-1) 原著論文発表

●論文詳細情報

1. 扇谷豪, 三木洋平, 朴泰祐, 森正夫, 中里直人, “計算宇宙物理のためのGPUクラスタ向け並列 Tree Code の開発と性能評価”, 情報処理学会 ACS 論文誌 (採択決定)
2. A. Akagic, H. Amano, "High-Speed Non-Adaptable and Fully-Adaptable CRC Accelerators," IEICE Trans. Inf. & Syst. "accepted"
3. Y. Shigeta, T. Inui, T. Baba, K. Okuno, H. Kuwabara, R. Kishi, M. Nakano, "Quantal cumulant mechanics and dynamics for multidimensional quantum many-body clusters", International Journal of Quantum Chemistry, vol. 113, No. 3, pp.348-355, 2013 (DOI: 10.1002/qua.24052).
4. T. Matsui, Y. Kitagawa, M. Okumura, Y. Shigeta, S. Sakaki, "Consistent scheme for computing standard hydrogen electrode and redox potentials", Journal of Computational Chemistry, vol. 34, No. 1, pp.21-26. 2013 (DOI:10.1002/jcc.23100)
5. Y. Takano, O. Okuyama, Y. Shigeta, H. Nakamura, "Density functional studies of the structural variety of the Cu₂S₂ core of the Cu-A site", International Journal of Quantum Chemistry, vol. 112, No. 24, pp. 3756-3762, 2012(DOI: 10.1002/qua.24214).
6. M. Shoji, K. Hanaoka, A. Sato, D. Kondo, M. Yang, K. Kamiya, K. Shiraishi, "Calculation of the Electron Transfer Coupling Matrix Element in Diabatic Reactions", International Journal of Quantum Chemistry, Vol. 133, No. 3, 342-347, 2013 (DOI: 10.1002/qua.24074)
7. S. Le Roux, A. Bouzid, M. Boero and C. Massobrio, "Structural properties of glassy Ge₂Se₃ from first-principles molecular dynamics", Physical Review B, Vol. 86, No. 22, 224201 (12 pages), 2012 (10.1103/PhysRevB.86.224201)
8. C. Massobrio, D. Mbongo Djimbi, M. Matsubara, R. Scipioni and M. Boero, "Stability of Ge₁₂C₄₈ and Ge₂₀C₄₀ heterofullerenes: a first principles molecular dynamics study", Chemical Physics Letters, Vol. 556, No. 2, 163, 2013 (10.1016/j.cplett.2012.11.033)
9. A. Akagic, H. Amano, "A Study of Adaptive co-processors for Cyclic Redundancy Checks on an FPGA," Proc. of International Conference of Field Programmable Technology, pp.119-124. Dec. 2012. 978-1-4673-2845-6/12 IEEE
10. T. Miyajima, D. Thomas, H. Amano, "A Domain Specific Language and Toolchain for OpenCV Runtime Binary Acceleration using GPU", Proc. of Third International Conference on Networking and Computing, pp.175-181, Dec.2012,

10.1109/ICNC.2012.34

11. 扇谷豪, 三木洋平, 朴泰祐, 森正夫, 中里直人, "重力多体系用 Tree Code の並列 GPU 化による計算加速", 2013 年ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2013 論文集, 2013.
12. T. Hanawa, Y. Kodama, T. Boku, M. Sato, "Tightly Coupled Accelerators Architecture for Minimizing Communication Latency among Accelerators", Proc. of 3rd Int. Workshop on Accelerators and Hybrid Exascale Systems (conj. with IPDPS2013) (accepted)
13. 塙敏博, 児玉祐悦, 朴泰祐, 佐藤三久, "Tightly Coupled Accelerators アーキテクチャに基づく GPU クラスターの構築", SACSIS2013 論文集(採択決定)