

jh150008-IS01

Xeon Phi・ベクトル計算機への FDTD コードと電磁流体コードの最適化手法の研究

深沢 圭一郎（京都大学）

概要 Xeon Phi は CPU 動作コードを変更すること無く動作可能であるが、そのままでは Xeon Phi の性能は全く出ないことがわかっている。また、SX-ACE は今までのベクトル CPU とは異なり、小さなベクトルチップからなる 4 ベクトルコア＋スカラコアの CPU となっており、最適化手法が確立されているとは言えない。そこで本研究課題では、電磁場解析を行う FDTD と宇宙プラズマを解く MHD コードを利用して、Xeon Phi システムと SX-ACE に向けた最適化を行い、その最適化手法をまとめることを目的とする。

1. 共同研究に関する情報

(1) 共同研究を実施した拠点名

北海道大学 情報基盤センター
東北大学 サイバーサイエンスセンター
東京大学 情報基盤センター
京都大学 学術情報メディアセンター
九州大学 情報基盤研究開発センター
上記の拠点所属の研究者と共同研究を行い、東北大学、京都大学、九州大学の計算機資源を利用した。

(2) 共同研究分野

(3) 超大規模情報システム関連研究分野

参加研究者の役割分担

- ・岩下 武史（北海道大学）：FDTD の XC30（Xeon Phi）での性能評価
- ・江川 隆輔（東北大学）：東北大学 SX-ACE を利用した MHD シミュレーションの性能評価と最適化
- ・大島 聡、片桐 孝洋（東京大学）：Xeon Phi を利用した FDTD 最適化支援
- ・南里 豪志（九州大学）：九州大学 HA8000 を利用した MHD シミュレーションの並列計算実行確認
- ・深沢 圭一郎（京都大学）：京都大学 XC30（Xeon Phi）、東北大学 SX-ACE を利用した MHD シミュレーションの性能評価・最適化・実行

2. 研究の目的と意義

現在の Top500 で 1 位となっている天河 2 号が Xeon と Xeon Phi で構成されていることに代表されるようにスーパーコンピュータシステムに何らかの加速器が組み込まれていることは一般的である。これらの加速器はその性能の高さから、システムの性能のほとんどを占めることが多い。

このような中、Xeon Phi は Native 利用においては CPU を利用することなく、独立した計算ノードとして利用することができ、さらに次期 Xeon Phi では bootable 対応となり、CPU 無しで、Xeon Phi だけで構成された計算機システムの登場が予定されている。Xeon Phi は多数の x86 系計算コアからなり、一般に CPU で動作しているアプリケーションであれば、コードの変更無く、動作可能である。しかしながら、そのままでは実行効率が極めて低く、Xeon Phi の性能は全く出ないことがわかっている。この原因はいくつか考えられるが、SIMD 利用効率が低いことが大きな問題の一つである。このように動作はするが性能が出ないために Xeon Phi を活用しているアプリケーションはほとんど無い。またポスト T2K システムでも Xeon Phi のようなメニーコアシステムが採用される計画で有り、Xeon Phi のようなメニーコアシステムを活用できるアプリケーションの準備は非常に

重要である。

また最新のベクトル計算機である SX-ACE も運用が始まっており、さらに次期ベクトル機として Aurora と呼ばれるシステムの開発も進められている。SX-ACE では今までのベクトル CPU とは異なり、小さなベクトルチップからなる 4 ベクトルコア＋スカラコアの CPU となっており、加速器付き CPU システムと構成が似ている。一般的にベクトル計算機はメモリバンド幅が大きいことが特徴であり、同構成の Xeon Phi や GPU と最適化の共通性がある可能性も考えられる。

そこで本研究課題では、電磁場解析アプリケーションである FDTD と電磁流体力学シミュレーションコードの MHD コードを利用して、Xeon Phi システムと SX-ACE に向けた最適化を行い、その最適化手法をまとめることを目的としている。

前述のように Xeon Phi はすでに複数の計算機センターで運用が始まっており、今後はポスト T2K も含め、さらに増えると思われる。アプリケーション開発者、利用者にとって、GPU と異なり、コードの変更無しで利用できる MIC (native 利用時) は魅力的で有り、利用者は多いと思われるが、同時に性能が出ない問題が増大すると予想される。そのような中で、実アプリケーションで利用頻度の高い流体系計算、電磁場解析コードにおいて、MIC 向け最適化手法を研究し、情報を計算機センターに蓄積することは、今後のセンターの役割からも非常に意義がある。スーパーコンピュータを利用して計算を行う研究者の大部分は大規模の計算が正確で早く解くことが出来れば、どのシステムで良いという考えである。その考えからも新型ベクトル機の最適化手法を研究することは近年ベクトル機を利用したことが無いユーザが多いことも考えると意義がある。特に前述の

ように CPU 構成の変わったベクトル機において、ベクトル機で利用頻度の高い流体系、電磁解析系アプリケーションで最適化を行うことは利用価値の高い情報提供につながると考えられる。

3. 当拠点公募型共同研究として実施した意義

本研究課題において、Xeon Phi と新型ベクトル機に対する実アプリケーションの最適化を行うことは、このアプリケーションを利用する研究者だけでなく、その最適化情報はアプリケーション開発者に広く有用な情報となる。

Xeon Phi は現状の HPCI-JHPCN において 3 計算機システムが利用可能である一方、実アプリケーションでの効率的な利用例がほとんど無い。今後 Xeon Phi から成る計算機システムが増えることを考えると、JHPCN に含まれる計算機センターと共同でアプリケーションの Xeon Phi への最適化を行うことは技術蓄積・公開の面からも非常に有用であると考えられる。

ベクトル機は昔から続いてきたロングパイプライン、高 B/F 値といった特徴が変化し、コア数の増加、ADB と呼ばれるメモリキャッシュ技術やスカラコアの導入が行われており、きれいなメモリアクセスと長い計算ループがあれば性能が出るという時代では無くなっている。このため、今までのベクトル最適化技術にスカラ的側面から最適化を行うことが重要と考えられる。このような状況においては、異なる計算機システムを運用する計算機センター研究者が共同研究することが、新しい知見を効率的に発見していくことの出来る仕組みの一つと考えられる。

4. 前年度までに得られた研究成果の概要

新規課題のため無し

5. 今年度の研究成果の詳細

まず、FDTD コードを、京都大学の計算機システムを利用し、Xeon Phi と Xeon 上で実行し、性能評価を行った。Xeon は Sandy Bridge (SB) と Haswell (HW) アーキテクチャの計算機システムを利用した。図 1、2 にそれぞれの結果を示す。計算サイズは $500 \times 500 \times 500$ に固定し、1 ノードを用いて計算し、スレッド数を変更させている。Xeon の場合は 2 CPU ソケットシステムのため、SB Xeon で 16 スレッドまで、HW Xeon で 28 スレッドまで利用している。SB Xeon では 2 から 4 スレッド、HW Xeon では 8 から 14 スレッドの間で性能向上があまりないが、利用スレッド (コア) の増加によるスレッドあたりのメモリ帯域が減少する影響と考えられる。その後は性能がスケールしている。最大性能はそれぞれ、約 13.6 GFlops (SB)、24.3 GFlops となっている。同じコードを Xeon Phi (Native 利用) 上でコアあたり 2 スレッドと 4 スレッドを利用して実行した結果、常に 2 スレッド/コアの性能が高くなった。それぞれ 48 コア (96 スレッド、192 スレッド) までは性能がスケールしているが、60 コアを利用した場合には性能が下がっている。最大の性能は 2 スレッドで 48 コア (96 スレッド) 利用時で、7.7 GFlops となった。1 CPU 辺りで Xeon と Xeon Phi を比べると、Xeon Phi が SB Xeon よりは性能が出ており、HW Xeon より性能が出ていないという結果となっている。

Xeon Phi の性能が出ていない原因を調べたところ、FDTD コード内で媒質定数に関連した部分が間接参照になっており、この部分で SIMD がうまく効いていない状況であった。媒質定数を全格子点分の要素を持つ配列と変

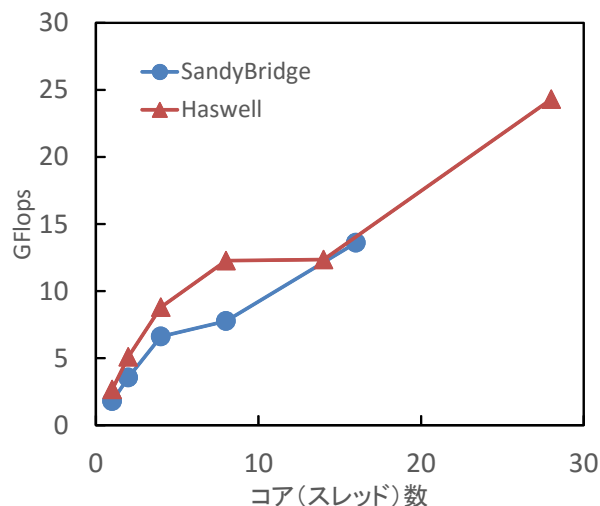


図 1 Xeon 上での FDTD コードの性能

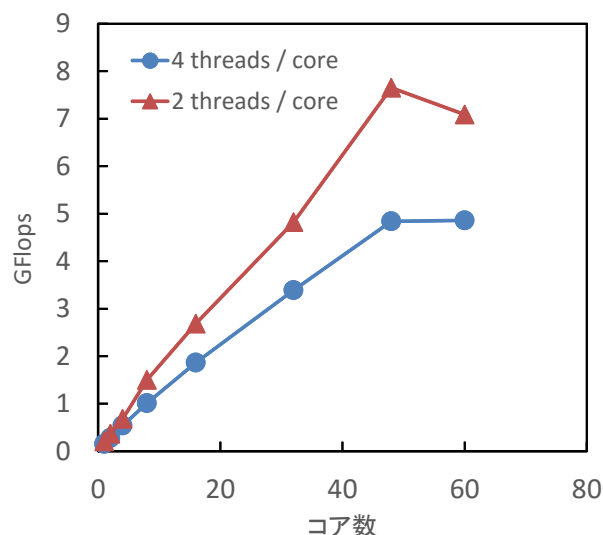


図 2 Xeon Phi 上での FDTD コードの性能

更したが、あまり性能改善は無かった。これは、変更により SIMD 自体は効くようになったが、ストリームの本数が増加し、またデータ転送量が増えるため、全体の性能としては改善が見られていない。これらの結果から、コードの構成をかなり変更する最適化を行わなければ、性能向上は難しいということが分かっている。

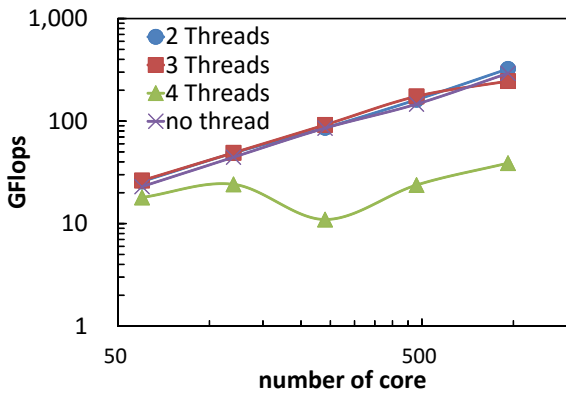


図 3 Xeon Phi 上での MHD コードの性能

MHD コードでは、Xeon Phi と SX-ACE を用いて性能評価を行った。MHD の単体 Xeon Phi 利用では、これまでの研究から、最適な領域分割法の種類の調査や Flat MPI では性能が出ないことなどが分かっている。京都大学の Xeon Phi システムでは Native 利用で複数ノードを利用できるため、複数ノードを利用し、スレッド数を変化させ性能を調べた。図 3 に MHD コードで 3 次元領域分割を行い、4 ノード (4 個の Xeon Phi) までを利用した結果を示す。ここでは、よく言われるように、プロセスあたりに 4thread を利用した場合が、非常に悪い性能になり、その他ではそこまで差が出ていない。Flat MPI の性能がこれまでの研究と異なり性能が高いが、評価システムが異なるため、詳細な理由は分かっていないが、スレッド利用時の性能が余り出でおらず、相対的に差が見えなくなっていると考えられる。このため、スレッド性能を上げる最適化が必要となっている。

次に、SX-ACE を利用し MHD コードを実行した際の結果を図 4 に示す。3 種類の領域分割とキャッシュヒット効率を考慮した 3 次元領域分割の 4 種類を評価した。これまでのベクトル機と同様にベクトル長が長い場合の性能が出ており、1 次元領域分割の性能が最も高かった。一方で、これまで性能が低くなっていたキャッシュヒット効率を考慮した場合 (type B) で、それほど性能が他に比べて

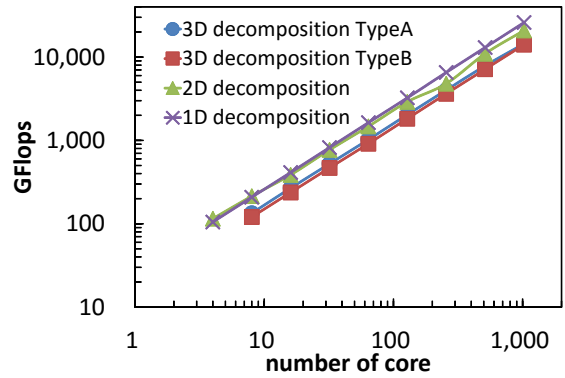


図 4 SX-ACE 上での MHD コードの性能

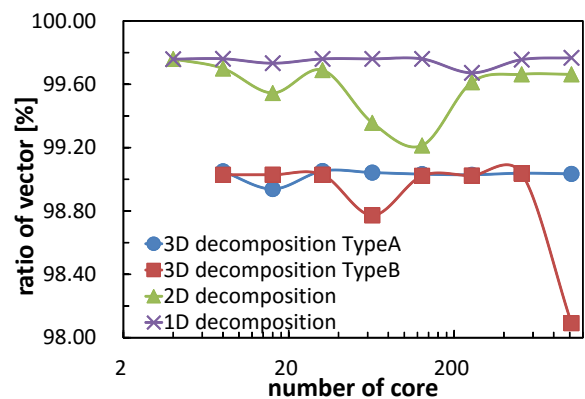


図 5 MHD コードのベクトル化率

も悪くなっていないかった。

図 5 にベクトル化率を図 6 に ADB 利用率を示す。1 次元と 2 次元領域分割でベクトル化率が高いことが分かり、3 次元領域分割では少し低くなっている。ADB (Assignable Data Buffer) はコアに付属のバッファメモリであり、キャッシュのような働きをする。SX-ACE では SX-9 に比べて容量の強化が行われ、4 倍のサイズ (1MB) になっている。3 次元領域分割 Type B において、ADB の利用率が高いことが図 6 からわかり、ADB が性能低下を防いでいると考えられる。図 7 にスケーラビリティを示す。2 次元領域分割の性能劣化は複数回計測でも現れているが、原因は不明である。その他をしてみると、この評価は weak スケーリングであるが、スケーラビリティが下がっていることが分かる。SX-ACE を利用した場合、ノード単体性能が高い一方で並列

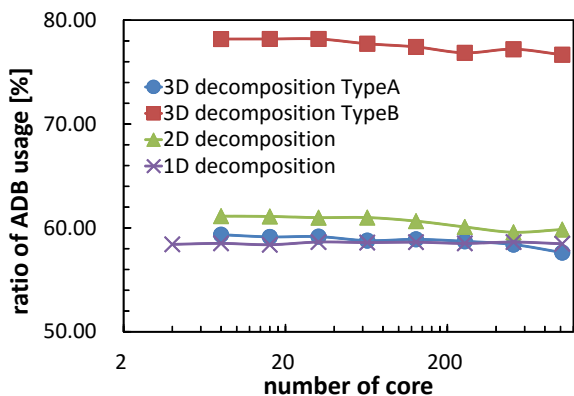


図 6 MHD コードの ADB 利用率

化性能が良くない結果となっている。

この SX-ACE における MHD コードの性能評価に対して、最適化を行った。最適化を行うべき点は、3次元領域分割の性能向上と並列化性能の向上である。並列度が上がっている現在では、ベクトル機であっても1次元領域分割は実用上利用しづらい。そこで、1次元領域分割で性能が高いという結果から、配列 $f(x, y, z)$ において、 x を長くし、 y, z を変更し、それに伴い、 y, z の領域分割数を変更した。表 1 にその結果を載せる。ここでは通信要素数と実際にかかった時間も計測している。まず、 y, z とともに極端に小さい場合は余り性能が出ておらず、100 程度が最適なサイズと考えられる。また、この場合が最も通信要素が少なく、実際に通信にかかった時間も短い。3次元領域

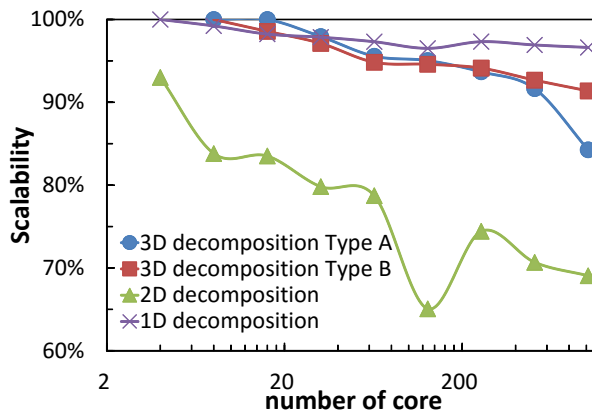


図 7 MHD コードの並列化性能

分割での並列化性能の向上の面からもこの $y = z$ の設定が最適だった。

次に x 方向のサイズ変化に伴う性能を調べた。図 8 に x 方向のサイズと実行効率の変化を示す。これから分かるように、 x 方向のサイズが大きいほど性能が出る傾向がある。これは、1次元領域分割の性能が高い結果と同じ傾向であり、ベクトル機の特徴と言える。ただし、そのサイズが 2 のべき乗の場合、バンク競合により極端に性能が劣化することが見えている。競合を避ける場合、その競合するサイズより -1 した場合に性能が最も出ている。ここでは最大で 44.7 % の実行効率が得られた。また、 y, z を 100 から 200 にすることで性能が少し上がることも確認された（実行効率

表 1. 計算サイズ、分割数を振った場合の性能測定結果

x	y	z	npex	npey	npez	通信要素数				通信処理		GFLOPS/proc (効率:%)
						x*y	y*z	z*x	Total	[sec]	[%]	
600	25	400	2	8	8	15,000	10,000	240,000	265,000	5.54	25.8	24.01(37.5%)
600	25	400	2	4	16					5.42	25.3	24.09(37.6%)
600	50	200	2	8	8	30,000	10,000	120,000	160,000	4.93	23.6	24.86(38.8%)
600	50	200	2	4	16					4.88	23.5	24.89(38.9%)
600	100	100	2	8	8	60,000	10,000	60,000	130,000	4.05	19.2	24.71(38.6%)
600	100	100	2	4	16					4.14	19.7	24.78(38.7%)
600	200	50	2	8	8	120,000	10,000	30,000	160,000	5.12	23.7	23.95(37.4%)
600	200	50	2	4	16					4.95	23.0	24.14(37.7%)
600	400	25	2	8	8	240,000	10,000	15,000	265,000	7.26	30.4	21.46(33.5%)
600	400	25	2	4	16					7.17	30.0	21.61(33.8%)

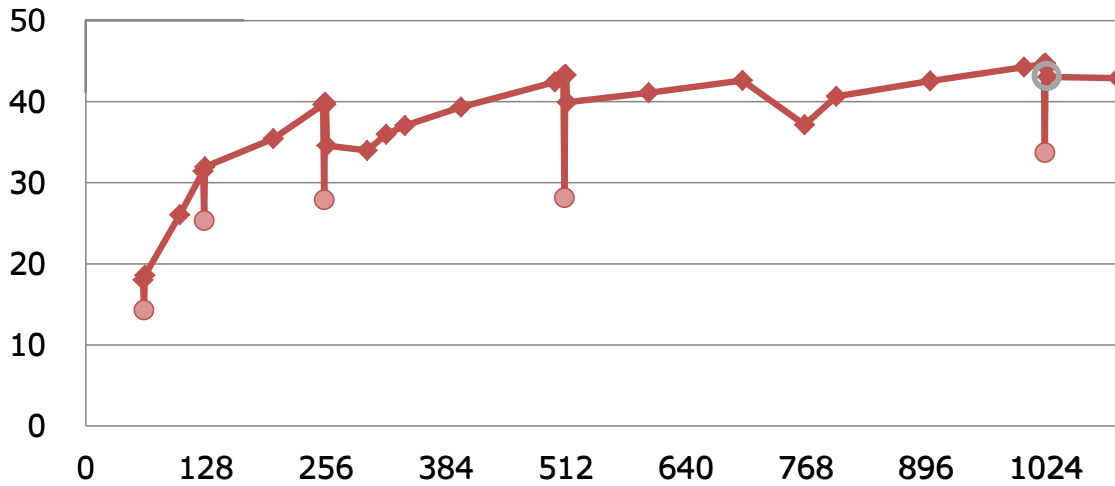


図 8 x のサイズ (ループ長) を振った場合の性能測定

45.2%)

この条件で並列化性能を再度計測してみると、図 9 のようになった。ここでは、プロセス当たりの計算サイズを(1023, 200, 200)として、1024 プロセスまで測定している。今まで並列化性能が下がっていたが、並列化効率 99.997% を達成することができた。

更なる最適化のため、MHD シミュレーションコードのメインループ解析を行った。コード上では、B/F 値が 3.59 となっているが、ADB を考慮した実際の実行時の B/F 値は 0.53 となっており、実効性能は演算ネックとなっていた (SX-ACE の B/F 値は 1)。MHD コードのメインループ内では、差分計算に伴う加算と乗算が多数出てくるが、それらの命令数の差が大きく、演算バランスが良くない。これにより理論性能がうまく出ていないことがわかり、冗長演算を共通化することで性能が向上することが確認できた。しかしながら、4%程度の性能向上で有り、この最適化によりコードの保守性が著しく損なわれるため、一般的には薦められない結果となった。

6. 今年度の進捗状況と今後の展望

本年度は、FDTD コードにおいては Xeon Phi での性能評価、最適化を研究計画としている。MHD コードは Xeon Phi と SX-ACE

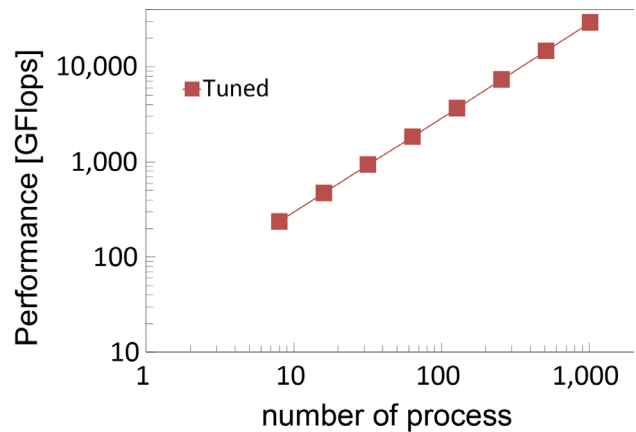


図 9 SX-ACE に最適化した MHD コードの性能

での性能評価、最適化を計画している。FDTD コードは京都大学 XC30 の Xeon Phi を利用し、性能評価を完了している。Xeon Phi で有効なスレッド数変化やメモリアクセス系のコンパイルオプションの最適化調査は行ったが、目立った性能向上は出ていない状況である。問題点がどこにあるのかを調査し、その問題を解決することが今後の課題となっている。

MHD コードは Xeon Phi での性能評価を実施し、こちらもスレッド数、コンパイルオプションの調査を行った。こちらのコードでは比較的的性能向上が見られている。Xeon Phi はスレッド数が大きいということを考慮した、ループの融合やストリーム数の変更など最適化を加えているが、目立った性能向上は

見えていない。また、SX-ACE での性能評価を行い、東北大学の磯部共同研究部門研究員の協力の下最適化を行った。今後は、この最適化結果に、Xeon Phi において実施した各種最適化を適用し、Xeon Phi と SX-ACE それぞれで効果的な最適化が、お互いのシステムでどのような効果を出すか調査する計画である（継続課題では採択されていないため、別途実施する予定）。

また、発表していない今年度の成果がまとめ次第、国内外会議での発表などを積極的に行っていく予定である。

計算機資源は、性能評価と最適化が主な利用方法のため、前半では利用率が低い資源（京都大学）もあったが、最終的には、東北大学、京都大学、九州大学ともに順調に計算機資源を利用できた。

7. 研究成果リスト

(1) 学術論文

1. Umeda, T., and K. Fukazawa, A high-resolution global Vlasov simulation of a small dielectric body with a weak intrinsic magnetic field on the K computer, *Earth, Planets and Space*, Vol.67, No.1, 49 (8pp.), 2015.

(2) 国際会議プロシーディングス

該当プロシーディングス無し。

(3) 国際会議発表

1. (Invited talk) K. Fukazawa, and R. J. Walker, "An MHD Simulation of the Dynamics of the Kronian Magnetosphere Driven by Solar Wind Observations", AOGS 12th Annual Meeting, Singapore, 2015.
2. K. Fukazawa, "An MHD Simulation of Global Planetary Magnetosphere Using K-Computer", AOGS 12th Annual Meeting, Singapore, 2015.

3. K. Fukazawa, R. Egawa, Y. Isobe, and I. Miyoshi, "Performance Evaluation of MHD Simulation Code on SX-ACE and FX100", poster in HPDC2016, Kyoto, 2016.

(4) 国内会議発表

1. 深沢圭一郎, "惑星磁気圏超並列高効率 MHD シミュレーションの開発", 名古屋大学 HPC 計算科学連携研究プロジェクトシンポジウム, 名古屋, 2015
2. K. Fukazawa, Yoshiyuki MORIE, Takeshi SOGA, Toshiya TAKAMI, and Takeshi NANRI, "Performance Evaluation of MHD Simulation Adopting a Specific Thread for Halo Region", Workshop on Information Technology, Applied Mathematics and Education, 津和野, 2015
3. 深沢圭一郎, 森江善之, 曾我武史, 高見利也, 南里豪志, "エクサスケールコンピューティングに向けた Halo スレッドの電磁流体シミュレーションに対する効果", 第 151 回 HPC 研究会, 那覇, 2015

- (5) その他（特許, プレス発表, 著書等）
特になし。