

CPU 消費電力制限印加時における電磁流体シミュレーションコードに対する配列構造最適化の評価

深沢圭一郎¹ 南里豪志²

概要: 近年スーパーコンピュータセンターやデータセンターにおいて、電源容量がその設計における拘束条件となっている。また、夏場など電力需要が高まる時期には、それらセンターで計算機の縮退運用が求められることも珍しくはない。このような状況において、消費電力を効果的に削減できる手法として、CPU 消費電力制限がある。CPU 消費電力制限には、周波数制限や利用コア数削減が考えられるが、よく利用される計算性能最適化手法がこの制限の影響をどのように受けるのかは明確になっていない。そこで、電磁流体シミュレーションコードに対し配列構造を変化させる最適化を施し、CPU 消費電力制限を加えることで、計算性能がどのように変わるのかを評価した。CPU 消費電力制限を掛けることで、Flops 値が変化し、bandwidth/Flops も変化するため、配列構造最適化の効果が変化する結果が得られた。本研究では、これらの評価結果を示し、その振る舞いについて議論する。

キーワード: 消費電力, 電力キャッピング, MHD シミュレーション, 性能評価

Evaluation of array configuration optimization for MHD simulation under CPU power capping

KEIICHIRO FUKAZAWA^{†1} TAKESHI NANRI^{‡2}

Abstract: Recently the power capacity became to restrict the design of supercomputer and data centers. In addition, the requirement of degenerate operations to the centers is not unusual in the high electricity demand season like summer. The CPU power capping is the one of most efficient way to reduce the power consumption. There are mainly two ways to cap the CPU power consumption, limitation of CPU frequency and decrease of CPU core. However, it is not clear how those capping ways affect to the optimization of computation performance. To understand the influence of capping, the evaluation of computation performance with the array configuration optimization to the MHD (magnetohydrodynamic) simulation code has been conducted under the CPU power capping. From the results, the behavior of optimization effect is changed due to the variation of Bandwidth per Flops (B/F) value corresponding to the CPU power limitation. In this study the results of performance evaluation and the behavior of optimization effect are discussed.

Keywords: Power consumption, Power capping, MHD simulation, Performance evaluation

1. はじめに

2020年11月にTop500で1位になった富岳では約30MWの電力が使われており、2010年1位のスーパーコンピュータ(スパコン)の消費電力と比べて7.5倍になっている[1]。一方で、スパコンセンターの電源容量はそれほど上がっておらず、また、運用上の電力コストを考えると、電源容量の著しい増強は難しい状況にある。このような状況のため、現在新たにスパコンを開発、設置する上で、システムの消費電力が最大の制限となっている[2, 3]。2020年~2021年にはエクサフロップス級スパコンが開発されると言われていたが、Green500(2020年11月現在)1位の電力効率で富岳の消費電力分を利用して780 PFlops程度の性能となり、エクサフロップスには達しない[4]。このままの電力性能であれば、エクサフロップスパコンを達成するには、40~50MWの消費電力が必要となる。50MWとなれば電力コストが、京都大学学術情報メディアセンターの電力コストの

20倍以上となり、運用が難しくなることが想像される。

このような電力問題を解決するために、CPUだけでなくGPUや特定の計算専用アクセラレータといったハードウェア開発が行われている。特にAI専用アクセラレータは現在その用途の多さからも世界中で様々な開発が行われており、実用化されているアクセラレータは高い電力性能を示している[4]。更に、運用における省電力を考えると、ハードウェア自体の電力性能だけでなく、電力制御機構を利用することで実消費電力を削減できる。富岳に搭載されているCPUであるA64FXであれば、電力を制御するパワーノブ機構を備えており、これらをアプリケーションに活用することで省電力が期待できる[5]。また、ジョブの消費電力を予測することで、冷却ファシリティを効率的に運用し省電力化を実現する運用手法も提案されており[6]、CPUの製造ばらつきによる電力性能を考慮した省電力スケジューリング手法も提案されている[7, 8]。

¹ 京都大学・学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto University
² 九州大学 情報基盤研究開発センター

Research Institute for Information Technology, Kyushu University

一方でパソコンを利用するアプリケーションレベルでは、計算性能を最大化し早く計算を終えることが主目的であるため、省電力に関して関心があまりない。そのため、アプリケーション自体の省電力化の研究はほとんど行われていないが、CPUの消費電力を制限することによるアプリケーションの省電力化手法は研究されている。例えば、Adagio Runtime [9] は DVFS (Dynamic Voltage and Frequency Scaling) を利用し、計算性能をほとんど下げずに消費電力を削減している。Sandy Bridge 世代以降の Intel CPU で利用可能な RAPL (Running Average Power Limit) を用いたアプリケーション実行時における CPU と DRAM への電力供給バランスの最適化研究が行われている[10]。今後は、アプリケーション自体の省電力性についても研究が必要であり、アプリケーション開発者は自分のアプリケーションがどのような消費電力特性を持っているか理解し、どのような電力最適化が効果的か理解しておく必要がある。

そこで本研究では、九州大学情報基盤研究開発センターのスーパーコンピュータシステム ITO が搭載している Skylake 世代の Xeon において、電磁流体力学 (MagnetoHydroDynamic: MHD) シミュレーションコードを利用し、CPU 消費電力制限下における配列構造最適化の評価を行った。MHD シミュレーションコードは流体シミュレーションコードの 1 種であり、今回の研究結果は一般の流体計算にも適用できると考えられる。

本研究報告の構成は以下の通りである。第 2 章では、スーパーコンピュータシステム ITO について説明し、第 3 章では MHD シミュレーションコードについて説明をする。第 4 章で消費電力測定の結果を述べ議論し、最後に研究のまとめをする。

2. スーパーコンピュータシステム ITO

スーパーコンピュータシステム ITO は、2017 年度に九州大学情報基盤研究開発センターに導入された計算機システムであり、Skylake 世代の Xeon を搭載している。システム高性能詳細は表 1 の通りだが、多数の CPU 計算ノードが接続されたサブシステム A (2,000 ノード) と 1 ノード当たり 4GPU が搭載されたサブシステム B (128 ノード) があり、本研究ではシステム A (ITO-A) のみを利用した。Skylake 世代の Xeon は Sandy Bridge 世代以降の Xeon のため、RAPL が利用できる。そこで本研究では、Inadomi らが開発した RAPL を利用するインターフェースである RIC を利用し、消費電力の測定、CPU 消費電力の制限をかけた場合の電力性能の評価を行っている[11]。

3. MHD シミュレーションコード

宇宙空間は真空と思われているが、その 99% はプラズマで満たされている。プラズマとは電離した気体のことであり、帯電している電子とイオンが分かれて存在する状態で

表 1 ITO サブシステム A の諸元

Table 1 Subsystem A of ITO.

機種名 Fujitsu PRIMERGY CX2550/CX2560 M4	
CPU	Intel Xeon Gold 6154 (Skylake-SP)×2 /node
コア数	18 cores /CPU
周波数	3.0 GHz (Turbo 3.7 GHz)
理論性能	3,5 TFlops /node (DP)
メモリ	DDR4 192 GB /node
Bandwidth	255.9 GB/s /node
B/F	0.074
総ノード数	2,000 nodes
総理論性能	6.91 PFlops
ノード間接続	InfiniBand EDR 4x (100Gbps)

ある。宇宙空間、特に我々の暮らす太陽系においては太陽から太陽風と呼ばれるプラズマの風が常時吹き出しており、太陽系全体にそのプラズマが充満している。このようなプラズマの振る舞いを記述する方程式として Vlasov-Maxwell 方程式がある。これは、無衝突 Boltzmann 方程式と Maxwell 方程式から成る。Vlasov (無衝突 Boltzmann) 方程式は以下の形をとる。

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \frac{\partial f_s}{\partial \mathbf{r}} + \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = 0 \quad (1)$$

ここで \mathbf{E} , \mathbf{B} , \mathbf{r} と \mathbf{v} はそれぞれ電場、磁場、距離、速度を表す。また、 $f_s(\mathbf{r}, \mathbf{v}_s, t)$ は位置-速度位相空間における分布関数であり、 S はイオンや電子など種類を示す。 q_s は電荷を m_s は質量を表す。

しかしながら、Vlasov 方程式は多くの成分からなる非線形方程式であり、計算機システムを用いても解くことが非常に難しい。そこで、Vlasov 方程式のモーメントをとることで求められる電磁流体力学 (MHD) 方程式が、グローバルなプラズマ構造を調べるときには使用されている。MHD 方程式は以下のようなになる。

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= -\nabla \cdot (\mathbf{v}\rho) \\ \frac{\partial \mathbf{v}}{\partial t} &= -(\mathbf{v} \cdot \nabla) \mathbf{v} - \frac{1}{\rho} \nabla p + \frac{1}{\rho} \mathbf{J} \times \mathbf{B} \\ \frac{\partial p}{\partial t} &= -(\mathbf{v} \cdot \nabla) p - \gamma p \nabla \cdot \mathbf{v} \\ \frac{\partial \mathbf{B}}{\partial t} &= \nabla \times (\mathbf{v} \times \mathbf{B}) \end{aligned} \quad (2)$$

上から、連続の式、運動方程式、圧力変化の式 (エネルギーの式)、最後が磁場の誘導方程式となる。簡単に言えば、電磁場を考慮した流体力学方程式と呼べる。詳しい導出方法は参考文献を参照されたい[12]。

MHD 方程式を解く数値計算法としては、Modified Leap

Frog (MLF) 法[13, 14]という計算法を使用する。これは最初の1回を two step Lax-Wendroff 法で解き、続く $(l - 1)$ 回を Leap Frog 法で解き、その一連の手続きを繰り返す。 l の値は数値的に安定の範囲で大きい方が望ましいので、本手法で採用する2次精度の中心空間差分では、数値精度の線形計算と予備的シミュレーションから $l = 8$ に選んでいる。

4. 消費電力性能評価

前述のように Skylake 世代の Xeon では RAPL が利用できるため、本研究では RAPL を利用する RIC という電力測定・キャッピング関数を用いて ITO の電力測定を行った。RAPL 利用時には CPU のハードウェアレジスタを参照、変更する必要があるが、九州大学情報基盤研究開発センターの協力によりそれらのレジスタを取り扱うことができた。

本研究では CPU 消費電力制限の計算性能最適化手法への影響を調べるために、これまで MHD シミュレーションコードの基本的な最適化として用いている配列構造最適化に注目する。MHD シミュレーションでは、空間3次元 (x, y, z) と MHD 変数 $(m = 8)$ をまとめた4次元配列を計算に利用している。通常は (x, y, z, m) という SoA 形式を用いているが、AoS 形式 (m, x, y, z) を用いると性能が向上する CPU もある。これらはベクトル向けかキャッシュ最適化向けかに大別することができるが、近年はこれらを両立させる場合に計算性能が上がることもあり、 (x, y, m, z) や (x, m, y, z) という配列構造も利用している。そこで今回はこれら4つの配列構造最適化に対して、CPU 消費電力制限がどのような影響を与えるか調査する。消費電力制限手法としては、RAPL による CPU 消費電力制限と利用 CPU コア数変化による消費電力制限を考える。

4.1 RAPL による CPU 消費電力制限下での最適化特性

CPU 消費電力制限では RAPL で PKG (CPU) 利用可能電力を 60W~160W の範囲として、4つの配列構造 (x, y, z, m) , (x, y, m, z) , (x, m, y, z) , (m, x, y, z) における MHD シミュレーションの性能を測定した (以降の配列構造をそれぞれ、 $xyzm$, $xymz$, $xmyz$, $mxyz$ とする)。これまでの消費電力測定により、ITO-A において MHD シミュレーションコードでは 150W 程度の PKG 消費電力になることが分かっているため、実効的に消費電力制限が始まる 160W 以下の測定を行った。実行は RIC の利用制限により、1プロセス/1ノードとし、8 (2×2×2 の3次元領域分割) プロセス×36スレッドの並列実行としている。

4.1.1 配列構造 $xyzm$ の測定結果

表2に $xyzm$ の性能測定結果を示す。ここでは、消費電力制限値 (Capping), メモリバンド幅/CPU 演算性能 (B/F) 値, 計算時間 (E-time), CPU (PKG) 消費電力, DRAM 消費電力と CPU 周波数 (FREQ) を測定している。測定は10回行い、その平均値を測定値としている (以降の測定も同

表2 CPU 消費電力制限下における $xyzm$ の測定結果

Table 2 Results of $xyzm$ under CPU power capping.

Capping [W]	B/F	E-time [s]	PKG [W]	DRAM [W]	FREQ [GHz]
60	0.155	32.74	59.79	35.69	1.77
70	0.126	29.65	69.79	35.93	2.18
80	0.110	29.44	79.76	36.10	2.50
90	0.098	27.18	89.66	36.28	2.81
100	0.091	30.89	99.59	36.11	3.01
110	0.085	29.51	109.67	36.21	3.21
120	0.082	29.16	119.61	36.11	3.33
130	0.079	28.37	129.57	36.23	3.47
140	0.075	28.78	139.49	36.22	3.63
150	0.074	28.25	148.51	36.26	3.69
160	0.074	27.44	150.83	36.27	3.71

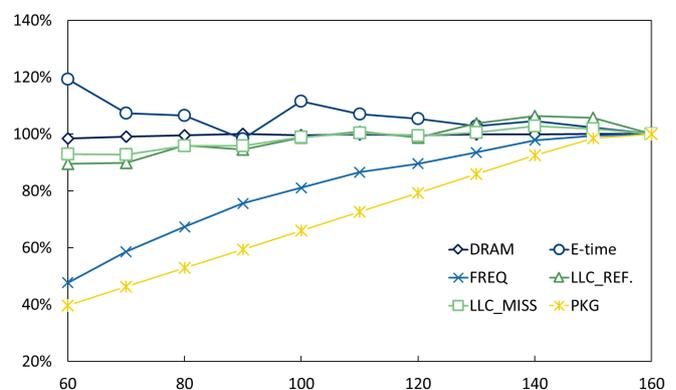


図1 CPU 消費電力制限による $xyzm$ の測定結果変化率
Figure 1 Variation of $xyzm$ results under CPU power capping.

様)。また、それら測定値に加え、LLC 参照数、ミス数の消費電力制限 160W 時との比を図1に示す。表2と図1から CPU の電力は消費電力制限に従い減少し、周波数も低くなっていることが分かる。DRAM 消費電力は変化無く、LLC の参照数、ミス数は両方共減少している。周波数は消費電力ほど下がっておらず、160~110W までは 0.1GHz/10W, 110~90W までは 0.2GHz/10W の割合で減少している。90~70W では約 0.3GHz, 70~60W では 0.5GHz の減少となっており、周波数の低下率は消費電力制限量に対して一律変化ではない。また、上記のように消費電力制限が強い場合に周波数低下率が大きくなる。

計算時間は一定の変化ではなく、ばらついており、想像していたほど低下が見えないが、120W 制限あたりから段々と遅くなっている。90W 制限で計算時間が制限無しと同じになるが、明確な理由は分かっていない。複数回の測定において同様の結果となっていることから、消費電力と計算性能のスイートスポットかもしれない。一般的に計算時間には周波数の変化が影響を与えやすいため、消費電力

表 3 CPU 消費電力制限下における xymz の測定結果

Table 3 Results of xymz under CPU power capping.

Capping [W]	B/F	E-time [s]	PKG [W]	DRAM [W]	FREQ [GHz]
60	0.164	32.44	59.83	34.48	1.67
70	0.132	31.28	69.80	34.57	2.07
80	0.114	30.32	79.76	34.71	2.40
90	0.101	28.51	89.71	34.94	2.72
100	0.093	28.06	99.67	34.79	2.94
110	0.087	28.32	109.59	34.76	3.15
120	0.082	29.82	119.60	34.76	3.36
130	0.079	28.23	129.56	34.76	3.46
140	0.077	26.77	139.53	34.86	3.57
150	0.074	29.30	148.98	34.61	3.69
160	0.075	26.93	151.64	35.00	3.68

表 4 CPU 消費電力制限下における xmyz の測定結果

Table 4 Results of xmyz under CPU power capping.

Capping [W]	B/F	E-time [s]	PKG [W]	DRAM [W]	FREQ [GHz]
60	0.169	31.51	59.81	34.07	1.62
70	0.133	30.39	69.77	34.17	2.06
80	0.115	27.28	79.76	34.28	2.39
90	0.100	28.63	89.71	34.23	2.73
100	0.093	29.64	99.68	34.07	2.93
110	0.087	28.32	109.63	34.10	3.16
120	0.082	28.46	119.60	34.10	3.34
130	0.080	28.63	129.54	34.04	3.44
140	0.077	27.01	139.44	34.19	3.58
150	0.075	28.44	148.07	34.07	3.66
160	0.075	27.32	149.48	34.52	3.64

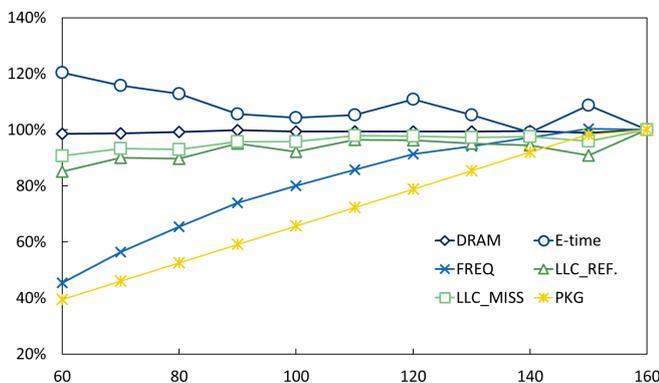


図 2 CPU 消費電力制限による xymz の測定結果変化率
Figure 2 Variation of xymz results under CPU power capping.

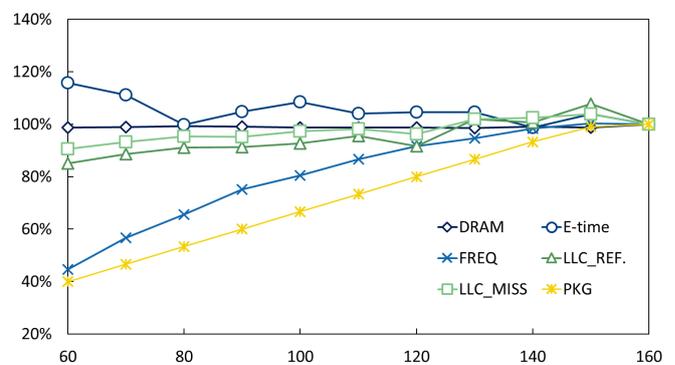


図 3 CPU 消費電力制限による xmyz の測定結果変化率
Figure 3 Variation of xmyz results under CPU power capping.

制限が大きくなるにつれて、計算時間は遅くなりやすい。今回の測定で、それほど遅くなっていない理由は、B/F 値が高消費電力制限時に高くなることが考えられる。160W 制限時と 60W 制限時では B/F 値が倍程度異なり、周波数の低下による性能劣化分を相殺している可能性がある。

4.1.2 配列構造 xymz の測定結果

次に xymz の測定結果を表 3 と図 2 に示す。xymz と同じように CPU 電力は、消費電力制限が大きくなるにつれて減少し、周波数も低くなっている。DRAM 消費電力は変化無く、LLC の参照率、ミス数は両方共減少している。この配列の場合でも、周波数は消費電力ほど低下せず、160~120W では 0.1GHz/10W、120~90W までは 0.2GHz/10W の割合で減少している。90~70W では約 0.3GHz、70~60W では 0.4GHz の減少となっており、周波数の低下率は消費電力制限量に対して一律の変化ではない。xymz に比べて周波数が大きく減り始める W 数が 10W 高く、同じ消費電力制限下でも xymz の周波数が 0.1GHz 低い状態で動作している。計算時間は、xymz 時に表れた 90W の特異点はなく、消費電力制限に応じて遅くなっている。

4.1.3 配列構造 xmyz の測定結果

これまでと同様に xmyz の結果を表 4 と図 3 に示す。CPU の電力と周波数変化は、他の配列の場合と同じ傾向になっている。詳しく見ると、周波数は、160~120W では 0.1GHz/10W、120~90W までは 0.2GHz/10W の割合で減少している。90~70W において約 0.35GHz、70~60W では 0.45GHz の減少率となっており、xymz や xmyz と比べ 60W 制限時の周波数が低くなっている。計算時間は xymz 同様にばらつきがあり、それほど低下が見えないが、130W 制限から遅くなっていることがわかる。80W 制限時に計算時間が制限無しと同じ程度になり、xymz と同様に特異点となっている可能性がある。xymz や xmyz に比べ、60W 制限時の周波数は低くなっているが、計算時間の増加率は xymz や xmyz より低くなっている（性能の低下率が低い）。

4.1.4 配列構造 mxyz の測定結果

mxyz の測定結果を表 5 と図 4 に示す。DRAM 消費電力は変化無く、LLC 参照数、ミス数は両方共減少している。CPU の消費電力が他の配列構造と比べて低くなっており、消費電力制限は 140W から効果が見える。その後は、消費

表 5 CPU 消費電力制限下における mxyz の測定結果

Table 5 Results of mxyz under CPU power capping.

Capping [W]	B/F	E-time [s]	PKG [W]	DRAM [W]	FREQ [GHz]
60	0.147	26.18	59.83	34.58	1.87
70	0.123	27.14	69.79	34.66	2.23
80	0.106	27.54	79.77	34.19	2.58
90	0.095	25.21	89.74	34.28	2.88
100	0.088	25.35	99.69	34.37	3.12
110	0.083	27.17	109.64	34.29	3.31
120	0.083	27.73	119.51	34.58	3.31
130	0.077	23.52	129.41	34.50	3.55
140	0.075	25.62	138.63	34.54	3.66
150	0.075	26.43	143.33	34.56	3.64
160	0.075	24.94	142.39	34.68	3.64

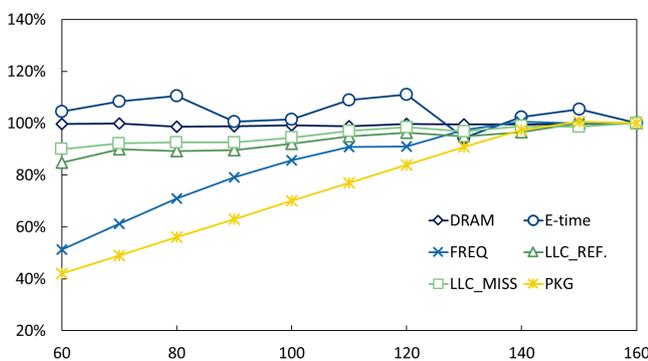


図 4 CPU 消費電力制限による mxyz の測定結果変化率
Figure 4 Variation of mxyz results under CPU power capping.

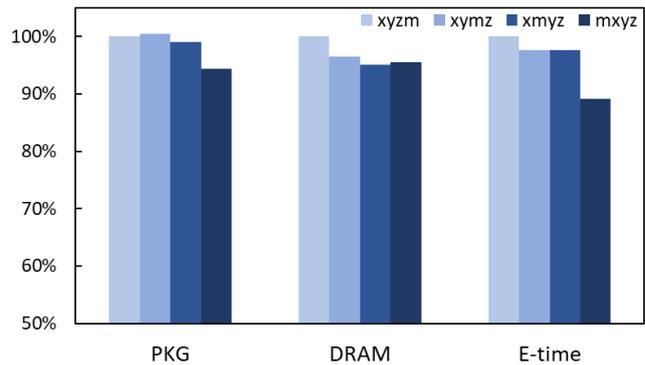
電力制限に従い減少していく。周波数は、他の配列構造の場合と異なり変化にばらつきがある。120W 制限時の周波数が少し低すぎるが、110W 制限までは他の測定結果と同様に 0.1GHz/10W の変化と考えられる。110~100W は 0.2GHz/10W で減少し、100~90W では 0.25GHz 程度、90~80W では 0.3GHz、80~60W では 0.35GHz の減少となっている。周波数の低下率は他の場合よりも低い結果となっており、60W 制限時の周波数が 1.87GHz と一番大きくなっている。

計算時間は、これまでと同様にばらつきが見える結果となっており、120W 制限から消費電力制限の影響が表れ始める。130, 100, 90W で計算時間が消費電力制限無しの場合と同程度となり、xyzm や xmyz で見られる特異点となっている。元々の CPU 消費電力が低いため、同じ消費電力制限時に他の配列構造よりも周波数低下が少なく、計算時間の増大率も最も低い値になっている。

4.1.5 配列構造同士の比較結果

4つの配列構造の消費電力制限無し (no Capping) と 60W 消費電力制限 (Maximum Capping) における CPU 消費電力

no Capping



Maximum Capping

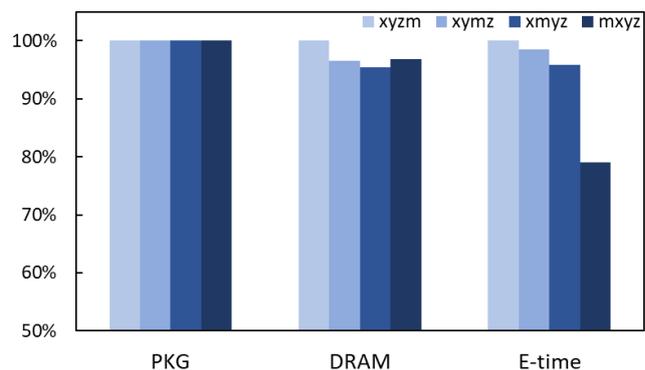


図 5 CPU 消費電力制限有り無しにおける各配列構造の測定結果比較

Figure 5 Comparison of array configuration optimization results with/without CPU power capping.

と DRAM 消費電力、計算時間を図 5 に示す。xyzm を 100% とした時のそれぞれ割合を示している。キャッピングが無い場合では、CPU 消費電力は mxyz が最も低く、次に xmyz となり、xymz は xyzm より少し高い結果となっている。DRAM 消費電力は xyzm だけが明らかに高く、最も低い配列構造は xmyz となっている。計算時間では mxyz が他の配列構造より明らかに低くなっている (計算性能が良い)。xymz と xmyz は同程度の計算時間となり xyzm より低くなっている。

60W 消費電力制限の場合では、RAPL による PKG 消費電力制限により CPU 消費電力はすべての配列構造で等しくなっている。DRAM の消費電力傾向は、消費電力制限が無い場合とほとんど同じだが、mxyz で xyzm と比した電力消費割合が大きくなっている。計算時間は大きな変化があり、xyzm に比べて mxyz の計算時間は消費電力制限無しよりさらに低くなっている。xymz と xmyz は消費電力制限無しでは同程度の計算時間だったが、xmyz の計算時間の方が短い結果となった。

mxyz は消費電力制限が無い場合に、CPU 消費電力が最も低く、その分が最大消費電力制限時に計算時間に振り返られて、mxyz の計算時間が相対的に低くなったと考えられ

る。割合は小さいが *xmyz* も同様である。このことから、最適化として *mxyz* を選んでおけば、消費電力制限時にも他の配列構造に比べて性能が高くなる事が分かる。一方で、消費電力制限が無いときに同程度の計算性能だった *xymz* と *xmyz* は、消費電力制限により性能に違いが出るため、消費電力制限有り無しで最適化の選択を変えるべきだということを示唆している。

4.2 利用 CPU コア数変化における最適化特性

CPU 消費電力は利用 CPU コア数を減らすことでも削減できる。そこで、CPU 消費電力制限下における配列構造最適化の結果と比較できるように、利用 CPU コア数変化時における配列構造最適化がどのように振る舞うのか調べる。本研究で利用している環境では、コア毎の電力が測定できないことや、CPU コアの C ステートを変更できないため、利用コア削減による CPU 消費電力の削減を測定できなかった。そのため、基本的には、CPU 消費電力キャッピング時と同様に計算性能の変化を調べる。第 1 近似的に CPU 消費電力は(利用コア数/利用可能最大コア数)の割合で変化すると考えておく。また、RAPL による CPU 消費電力制限時に変化していた CPU 周波数変化は、理論性能変化を意味するため、ここでは利用コア削減時の理論性能を記載する。電力性能測定実行において、動作の関係上 Flat MPI を用いて実行しており、CPU 消費電力制限時と計算時間に違いがある。実行は strong scaling であり、利用コア数によって計算サイズは変わらない (300×300×400 の一定)。

表 6 に *xyzm*, *xymz*, *xmyz*, *mxyz* の配列構造において、利用コア数を 36 コア (2CPU) から 2 コアまで変化させた場合の計算時間、理論性能、B/F を示す。また、各利用コアにおいて、*xyzm* の計算時間を 100%とした場合の配列構造間での性能比較を図 6 に示す。計算時間は、36 コアでは *mxyz* が最も良いが、利用コア数が減る毎に *xyzm* の性能が良くなっていくことが分かる。12 コア利用時にそれらの性能は逆転する。*xymz* と *xmyz* の性能変化はそれらと異なり、36 コア利用時に 2 番目に早かった *xmyz* は 2 コア利用時でも 2 番目に早い。*xymz* は 36 コア利用時の性能が最も悪かったが、2 コア時でも *mxyz* の次に性能が悪くなっている。これらの変化は B/F 値の改善が *mxyz* 以外には大きく効いているが、*mxyz* はキャッシュ利用効率が高い配列構造のため、その恩恵をあまり受けていないことが考えられる。つまりコードの実行時における (実際にメモリアクセスを行うのかキャッシュを参照できるのかを考慮した) の B/F 値が *xyzm* > *mxyz* というために起こる可能性が高い。

次に、理論性能 (Rpeak) の変化と計算時間の逆数として考えられる計算性能 (Rmax) の変化を比べる (並列化による性能向上率と同じ意味)。図 7 に 36 コア利用時を 100%とした計算性能の変化を示す。18 コア利用時は、すべての配列構造において理論性能と同様に 50%程度の性能になるが、16 コア利用時は理論性能に比べて性能がそれほど低下

表 6 利用コア数削減における計算時間

Table 6 Calculation time with decreasing CPU core usage.

Cores	Rpeak [GFlops]	B/F	E-time [s]			
			<i>xyzm</i>	<i>xymz</i>	<i>xmyz</i>	<i>mxyz</i>
2	192	0.666	106.31	124.39	116.71	139.78
4	384	0.333	58.41	67.81	61.69	74.72
8	768	0.167	38.46	39.09	35.07	41.10
12	1152	0.111	29.64	32.95	29.97	31.36
16	1536	0.083	27.36	27.88	25.88	26.02
18	1728	0.074	26.35	26.54	25.33	25.11
36	3456	0.074	13.29	13.61	12.80	12.74

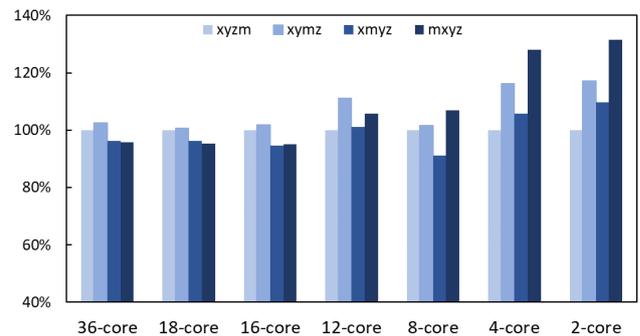


図 6 利用コア数削減時における計算時間の変化率比較
Figure 6 Comparison of calculation time with decreasing CPU core usage.

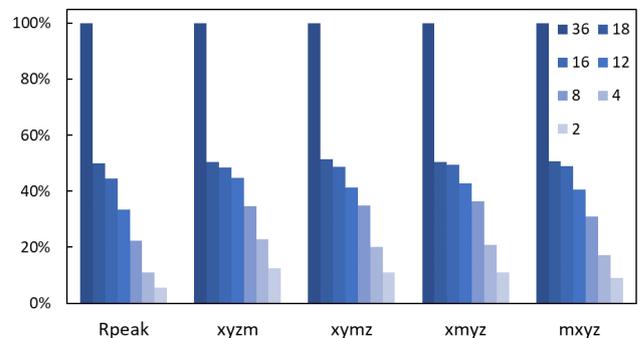


図 7 利用コア数削減における配列構造の計算性能変化率
Figure 7 Performance variation of array configuration decreasing CPU core usage.

しないことが分かる。12 コア利用時は、*xyzm* 以外は理論性能より性能変化率は高いが、これまでより性能の変化率が大きくなる。8 コア以降では、*mxyz* の性能が大きく下がっていくことが分かる。*xyzm* と比較すると、*mxyz* のコア数変化に対する性能の低下率変化が大きいことが良く分かる。

これまでの測定結果では、利用コア数削減は、CPU 消費電力制限より計算性能を大きく削減するため、省電力効果が高いが、実行時間増大に繋がる。ただし、2 コア程度の

利用コア数削減であれば、計算性能劣化はほとんど表れていないため、計算性能をある程度維持したまま省電力化に利用できる可能性がある。また、利用コア数削減に伴いB/F値も変化するため、最適配列構造が36コア利用時と12コア利用時以下では真逆になるという結果になった。消費電力制限時におけるB/F値変化も含めて考えると、10%程度のB/F値変化であれば計算性能に影響は少ないが、2倍変化するような状況では、最適化手法も変える必要があると考えられる。

5. まとめ

本研究では、九州大学のスーパーコンピュータシステムITOにおいてMHDシミュレーションコードを利用して、CPU消費電力制限下での配列構造最適化の効果を調べた。MHDシミュレーションにおいて、ベクトル化やキャッシュ利用といった面で配列構造が性能に大きく影響しているため、4つの配列構造 (x, y, z, m) , (x, y, m, z) , (x, m, y, z) , (m, x, y, z) を利用し、2つの消費電力削減方法を併用した際の最適化効果を調べた。RAPLによるCPU消費電力制限の場合では、配列構造によってCPU周波数の消費電力制限による低下の振る舞いに違いが見られ、計算性能の変化に影響を与えていた。また、計算時間が想定よりも大きく悪化することがなく、消費電力制限による理論的な低下分をB/F値の上昇分である程度相殺している可能性がある。消費電力制限が無い場合においてCPU消費電力が少ない配列構造がCPU消費電力制限時では、高い計算性能を出すことが分かった。

利用コア数削減による消費電力削減を考えた場合、利用コア数の数により、最適な配列構造が変わることが分かった。この場合も利用コア数削減に伴い、B/F値が変化するため、計算機の性質が変化し、最適化に影響を与えている。ただし、利用コア数削減は消費電力を大きく削減できると共に計算性能を大きく悪化させるため、注意が必要である。今回の結果であれば、18コア利用から16コア利用に削減する程度であれば、大きな計算性能削減なしに、消費電力をわずかに削減することが可能である。

CPU消費電力制限と利用コア数削減では、計算性能変化に違いがあった。消費電力を半分にするなど大きく減らす場合は、CPU消費電力制限をした方が性能に悪影響が少ない。ただし、MHDシミュレーションコードのような比較的高B/F値を持つアプリケーションには、B/F値変化が効いてくるため、アプリケーションの情報をうまく取得して、省電力コントロールをすることが望ましい。

謝辞 本研究は、九州大学情報基盤研究開発センター令和2/3年度先端的計算科学研究プロジェクトの支援による。

参考文献

- [1] The Top 500 site. (<https://www.top500.org/>)
- [2] P. M. Kogge, et al., Exa Scale Computing Study: Technology Challenges in Achieving Exascale Systems, in Exascale Computing Study Report, 2008. (http://users.ece.gatech.edu/mrichard/ExascaleComputingStudyReports/exascale_final_report_100208.pdf)
- [3] P. M. Kogge, and T. J. Dysart, "Using the TOP500 to trace and project technology and architecture trends," High Performance Computing, Networking, Storage and Analysis (SC), 2011 International Conference for, pp.1,11, 12-18 Nov. 2011.
- [4] The Green 500 Site. (<http://www.green500.org/>)
- [5] 児玉 祐悦, 小田嶋 哲哉, 有間 英志, 佐藤 三久, スーパーコンピュータ「富岳」における電力制御の評価, 研究報告ハイパフォーマンスコンピューティング (HPC), 2020-HPC-175 (25), 1-8, 2020.
- [6] 鈴木 成人, 平岡 美智子, 白石 崇, クリシュパ エンジ, 山本 拓司, 福田 裕幸, 松井 秀司, 藤崎 正英, 宇野 篤也, 高効率ファシリティアマネジメントを実現するジョブ電力予測手法の提案, 研究報告ハイパフォーマンスコンピューティング (HPC), 2019-HPC-172 (22), 1-12, 2019.
- [7] L. Li, K. Fukazawa, H. Nakashima, T. Nanri, A Node Level Performance/Power Efficiency Aware Resource Management Technique, 研究報告ハイパフォーマンスコンピューティング (HPC), 2018-HPC-166 (3), 1-7, 2018.
- [8] Jiacheng Zhou, Keiichiro Fukazawa, Hiroshi Nakashima, Energy Aware Scheduler of Single/Multi-node Jobs Exploiting Node Heterogeneity, 研究報告ハイパフォーマンスコンピューティング (HPC), 2021-HPC-178 (13), 1-12, 2021
- [9] B. Rountree, D. K. Lowenthal, B. de Supinski, M. Schulz, V. W. Freeh, and T. Bletsch, "Adagio: Making DVS practical for complex HPC applications", Proceedings of the 23rd international conference on Supercomputing, June 08-12, 2009, Yorktown Heights, NY, USA doi:10.1145/1542275.1542340.
- [10] K. Fukazawa, M. Ueda, Y. Inadomi, M. Aoyagi, T. Umeda, K. Inoue, "Performance Analysis of CPU and DRAM Power Constrained Systems with Magnetohydrodynamic Simulation Code", HPC2018, 2018.
- [11] Inadomi, Y., et al., "Analyzing and Mitigating the Impact of Manufacturing Variability in Power-Constrained Supercomputing", Technical Paper, SC'15, Austin (USA).
- [12] F. F. Chen, 1974. Introduction to Plasma Physics. Plenum Press, NY.
- [13] T. Ogino, R. J. Walker, M. Ashour-Abdalla, A global magnetohydrodynamic simulation of the magnetopause when the interplanetary magnetic field is northward, IEEE Trans. Plasma Sci.20, 817.828, 1992.
- [14] Fukazawa, K., T. Ogino, and R.J. Walker, "The Configuration and Dynamics of the Jovian Magnetosphere", J. Geophys. Res., 111, A10207, 2006.