

ハイブリッドメモリを搭載するシステムにおける データサイズを考慮した電力制御

有間 英志^{1,a)} 埴 敏博^{1,b)}

概要：将来の HPC システムにおいては、電力制約やメモリシステムの容量・バンド幅限界がシステムの性能向上を制限することが危惧されている。電力制約下においてシステムの総スループットを向上させるためには、各ジョブ/ノードに電力制約を与える技術(電力キャッピング)や、ボトルネックのコンポーネントに電力を優先的に割り当てる技術(電力シフティング)が不可欠である。一方で、メモリシステムの大容量化・高バンド幅化も重要な課題であり、ベンダーはこれらを両立するために、異なるメモリ技術によってメインメモリを構成するハイブリッドメモリをサポートし始めている(例: Intel 社製 Knights Landing)。しかし、この様な背景に反して、ハイブリッドメモリを搭載する実システム上での電力管理技術に関する研究は、十分になされていないのが現状である。そこで本稿では、その様なシステムにおける電力制御技術、特にアプリケーションの利用データサイズを考慮した電力制御技術を検討する。これは、ハイブリッドメモリでは、利用するデータサイズに応じて、その実効バンド幅が大きく変化する点に着目したものである。本稿では、ハイブリッドメモリを搭載する実システムにおいて、アプリケーションの利用データサイズを変化させつつ電力キャッピングや電力シフティングを行った結果を報告する。

Footprint-Aware Power Management on Hybrid Memory Based Systems

EISHI ARIMA^{1,a)} TOSHIHIRO HANAWA^{1,b)}

1. はじめに

昨今の HPC システムでは消費電力こそが主要な設計制約となっている。厳しい電力制約のもと高い性能を得るためには、システムの総電力効率を高める様な洗練された電力マネジメント手法が不可欠である。この様な背景のもと、電力キャッピング [22] — 各々のジョブ/ノードに対して電力制約を与える制御手法 — 及び、電力シフティング [11] — 各ジョブ/ノード内にて、より電力を必要としているコンポーネントに対して、電力を優先して割り当てる制御手法 — が有力であると考えられている。これまでには、電力予算を各ジョブやノードに対して割り当てる様々な電力スケジューリング手法が提案されてお

り [5], [9], [30], [32]、その様な割り当てられた電力を最大限有効利用するための電力シフティング手法もまた幾つか提案されている [3], [7], [13], [20], [23], [24]。

一方、メインメモリにおけるバンド幅及び容量制約もまた HPC システムにおいて主要な問題となりつつある。バンド幅を向上させるためには、メインメモリを HMC [6] や HBM [34] といった、最新の三次元積層 DRAM を用いて構築することが有力である。しかし、これらの技術では容量のスケラビリティに限界があり [19]、これは既存の DDR ベースの DRAM よりも深刻であることが知られている [17]。大容量化を達成するためには、近年注目を浴びているスケラビリティの高い NVRAM(Non-Volatile RAM, e.g. PRAM [21], ReRAM [1], STT-MRAM [2], [25]) を用いてメインメモリを構成することが有力である。しかし、これらのメモリ素子は DRAM よりも低速であり、性能の面で劣るという問題がある。この様な背景から、ベンダー

¹ 東京大学情報基盤センター
Information Technology Center, The University of Tokyo
a) arima@cc.u-tokyo.ac.jp
b) hanawa@cc.u-tokyo.ac.jp

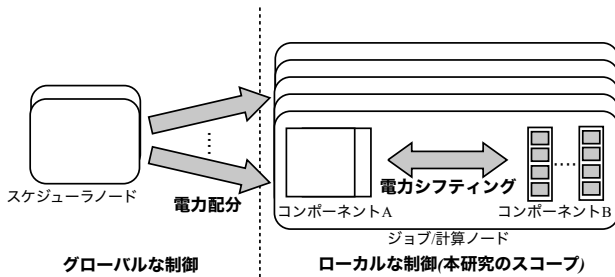


図 1 大規模システムにおける階層的電力制御

は徐々にハイブリッドメモリアーキテクチャに設計をシフトしつつある。これは、メインメモリを複数の異なるメモリ技術で構成し、その上でのデータ配置制御を最適化することで、高性能化と大容量化の両立を図るものである (e.g. 3D stacked DRAM + DDR-based DRAM [17] or DRAM + NVRAM [4], [35])。

このような状況を鑑みると、ハイブリッドメモリを搭載するシステムを対象とした電力マネジメント手法こそが、現在～将来にわたって必要な技術であると言える。本研究では、特にノード単位での電力制御に焦点を当て、与えられた電力予算のもと性能を最大化するための、データサイズを考慮した電力シフティングを提案する。これは、そのようなシステムにおいて、アプリケーションの利用データサイズをスケールさせると、メモリシステムの実行バンド幅が大幅に低下し、結果として、性能ボトルネックがメモリシステム側に移ることに着目したものである。従って、その場合には、電力割り当てもメモリシステム側にシフトさせる必要がある。本稿では、この有効性を確認するための予備的評価を行ったのでこれを報告する。

本論文の続きは次の様に構成される。まず次章では、背景及び関連研究について述べる。第3章では本研究の動機付け及びコンセプトについて述べる。第4章では評価方法及び評価結果について述べ、提案の有効性を定量的に示す。最後に第5章では本論文のまとめ及び今後の方針について述べる。

2. 背景及び関連研究

これまでに、データセンタやスーパーコンピュータ等大規模システム向けに様々な電力制御技術が提案されている。また、ハイブリッドメモリを搭載するシステム向けの電力制御技術に関してもこれまでに提案されている。本章では、本研究とこれらとの違いを明確化する。

2.1 大規模システムにおける電力制御

大規模システムの電力マネジメントは図1に示す様な階層的構造を取ることが多く、これまでの先行研究もグローバル制御 (電力を考慮したスケジューリング) 及び、ローカル制御 (電力シフティング) に分類できる。前者は、スケ

ジューラによるノード/ジョブへの電力配分最適化であり、後者は各ノード/ジョブにおける与えられた電力予算の有効活用であり、我々の研究は後者に属する。以下各々の詳細について述べる。

2.1.1 電力スケジューリング: グローバルな電力制御

昨今の大规模システムは深刻な電力制約を受けているため、各々のジョブ/ノードへ電力を分配するためのスケジューラベースの方法が提案されている。X. Fan らは大规模データセンタ向けにオーバープロビジョニングという概念を提案し、そこでの電力配分手法を提案している [10]。その後、T. Patki らは HPC システムにおけるオーバープロビジョニングの有効性の検証を行った [27]。これらの研究に基づき、様々なスケジューラベースの電力管理手法がこれまでに提案され、実装されてきた [5], [9], [30], [32]。これらの研究は大変有望であるものの、我々の研究は与えられた電力予算の中で各ノードの電力制御を最適化することに着目しているため、我々の研究とは直交するものである。

2.1.2 電力シフティング: ローカルな電力制御

電力シフティングという概念は先行研究 [11] にて登場し、これを可能にするため電力キャッピングが提案された [22]。それ以降、幾つかの電力シフティング手法がこれまでに提案されてきた。しかし、我々の研究はハイブリッドメモリを搭載するシステム上にて、アプリケーションのデータサイズを考慮しつつ、CPU とメモリシステムの電力を同時に最適化するという点で新しい。

幾つかの先行研究では、単一ノード内にて、プロセッサ (CPU または GPU) とメモリ間の電力シフティングに焦点を当てている [7], [11], [13], [28], [31]。これらの研究では、プロセッサとメモリの内どちらかが性能ボトルネックになっているのかを考慮して電力割り当てを最適化している。しかし、このような性能ボトルネックはハイブリッドメモリ上においてはデータサイズに大きく依存するにもかかわらず、この点に着目した研究は我々の知る限り本研究が初めてである。

その他にも様々なコンポーネント間での電力シフティングを行っている研究が存在する。R. Ge らは NUMA アーキテクチャ内の各 CPU 間の電力配分を考慮している [12]。CPU と GPU 間の電力シフティングを考慮した研究も存在する [3], [20]。その他の研究は CPU とネットワーク間の電力割り当て最適化を考慮している [23], [24]。最近の研究では、I/O を考慮した電力シフティングを提案しており、I/O インテンシブなフェーズにおいて CPU の電力を削減している [33]。しかし、これらの研究はメモリの電力に着目していないため、我々の研究とは直交している。

2.2 ハイブリッドメモリを搭載するシステムにおける電力制御

論文 [8] にて最初にハイブリッドメインメモリアーキテ

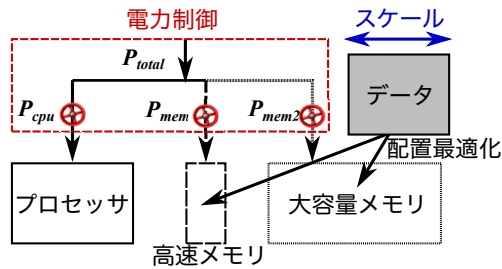


図 2 対象とするシステム

クチャが提案されて以来、様々な電力制御メカニズムが提案されてきた。H. Park らは DRAM 及び PRAM によって構成されたハイブリッドメインメモリについて、DRAM 領域をキャッシュとして利用し、cache-decay [18] を適用することで、リフレッシュ電力を削減する手法を提案している [26]。その他の研究では、DRAM-PRAM ハイブリッドメモリにおける、ページ単位のデータ配置制御を工夫することで、PRAM への書き込みエネルギーの削減を行っている [29] [37]。これらの研究では、ハイブリッドメモリシステム自体の電力削減を対象としているのに対し、我々の研究ではハイブリッドメモリを搭載するノードの電力制御を対象としているため、より包括的である。さらには、これらの研究ではシミュレータを用いた検証をしているのに対し、我々の研究では実システムを使って検証している点が異なる。

3. 動機付け及びコンセプト

本章では、対象とするシステム、本着想に至った経緯・動機付け、本研究のコンセプトの説明を行う。

3.1 問題設定

図 2 に、本研究が対象とするシステムを示す。プロセッサ、高速 (小容量) メモリ、大容量 (低速) メモリの電力制御が我々のスコープである。高速メモリと大容量メモリの間のデータ制御は、ソフトウェア (IMDT: Intel Memory Drive Technology [16] 等) もしくはハードウェア (KNL: Intel Knights Landing におけるキャッシュモード等) が自動的に行う場合を想定する。その様な環境において、様々なスケールのデータを使ってアプリケーションを動かす状況を想定する。

この様なシステムにおいて、スケジューラから当ノードに与えられた電力予算を P_{total} とし、これをプロセッサ/高速メモリ/大容量メモリの電力 $P_{cpu}/P_{mem}/P_{mem2}$ に上手く分配し、性能最適化を図る。具体的には、性能を $Perf()$ 、データサイズを C とおけば、以下の最適化問題に帰着される。ただし、 $P_{cpu}, P_{mem}, P_{mem2}$ は制御変数であるのに対し、データサイズ C はユーザ側で定められる given な変数であることに注意されたい。

```

1 #pragma omp parallel for simd
2 for (i = 0; i < N; i++) {
3     A[i] = A[i] * B[i] ... * B[i];
4 }

```

図 3 テストに用いたコード: 配列サイズ N を変化させることでデータサイズをスケールさせ、 $*B[i]$ の個数を変化させることで F/B 比をスケールさせている。

$$\max Perf(P_{cpu}, P_{mem}, P_{mem2}, C)$$

$$s.t. P_{cpu} + P_{mem} + P_{mem2} \leq P_{total}$$

3.2 動機付け

この様なシステムにおける電力割り当てを最適化することを目的とし、その制御手法の手がかりを得るため、我々は予備評価を実施した。具体的には、表 1 に示す KNL ベースのシステム及び、表 2 に示す SKL ベースのシステムを用いており、これらシステム上にて図 3 に示す様なコードを実行し、データサイズ、 F/B (Flops/Bytes) 比を変化させながら性能を計測した。

この評価結果を図 4 に示す。図では横軸は F/B 比を示しており、縦軸は性能 (GFLOPS) を示している。両軸ともに log スケールになっており、この曲線の概形はルーフライン曲線としてよく知られている [36]。すなわち、 F/B 比が高い場合にはプロセッサのスループット性能による制限を受け水平の線になるのに対し、 F/B 比が低い場合には、メモリシステムのバンド幅の制限を受けて斜線の概形になる。

表 1 KNL ベースのシステム

CPU Package	Xeon Phi x200 Processor 7210 (Knights Landing), 64 cores, 1.3GHz, In-package cluster mode: quadrant
Memory System	MCDRAM: 16GiB, 450GiB/s, DDR4: 96GiB, 90GiB/s Memory mode: flat, Data management: numactl -preferred=1
OS	Cent OS 7.2
Compiler	Intel C++/Fortran Compiler 17.0.4, Options: -O3 -qopenmp, -xMIC-AVX512

表 2 SKL ベースのシステム

CPU Package	Xeon Gold 6140 Processor (Skylake), 18 cores, 2.4GHz, TDP 140W x2sockets
Memory System	DRAM: 96GiB, 120GiB/s, TDP 70W, NVRAM: Intel Optane SSD P4800X, 375GB, 2.4GB/s(read), 2.0GB/s(write) x2, Data management: Intel Memory Drive Technology (IMDT) [16]
OS	Cent OS 7.4
Compiler	Intel C++/Fortran Compiler 17.0.4, Options: -O3 -qopenmp

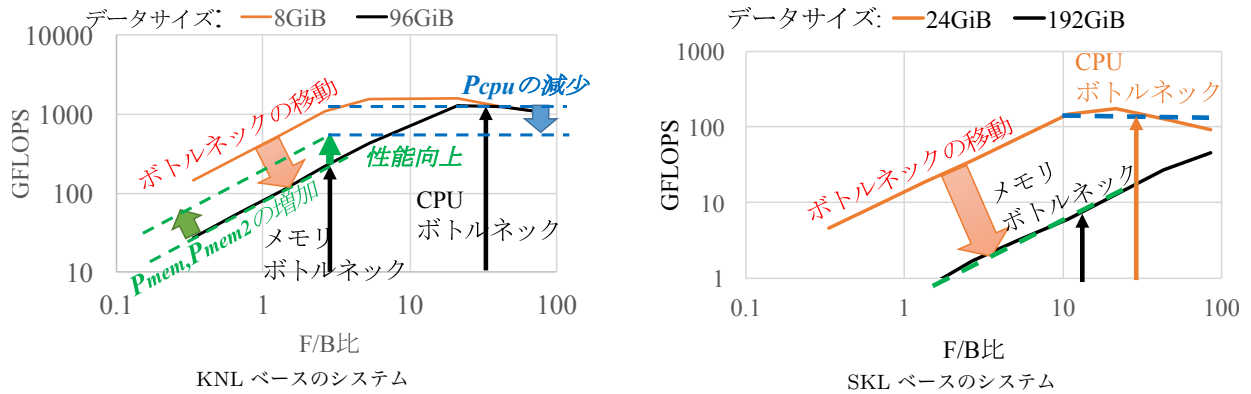


図4 ルーフライン曲線 [36] 及び提案コンセプト

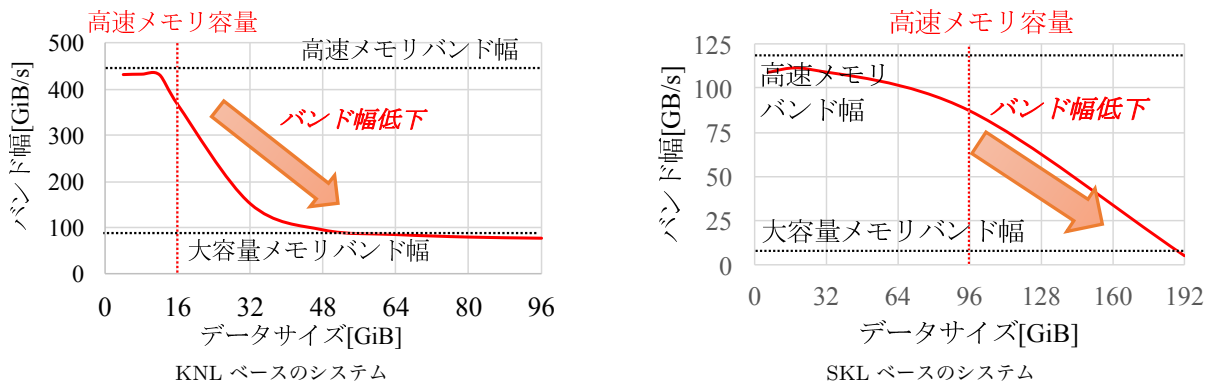


図5 実効バンド幅とデータサイズの関係 (F/B rate: 1/3)

図4の通り、両システムともにデータサイズを大きくすることで、メモリバンド幅律速の斜線部が大きく下側に平行移動しており、結果として多くのF/B比においてCPUからメモリシステム側にボトルネックが移動していることが分かる。この理由を図5に示す。図5の横軸はデータサイズを表しており、縦軸はメモリシステムの実効バンド幅を示している。グラフから分かる通り、データサイズが大きくなればなるほど、特に高速メモリの容量を超えると、実効バンド幅は大きく低下する。これは、データサイズが大きくなればなるほど、より頻繁に大容量(低速)メモリがアクセスされるようになるためである。この様に、より大きなデータを扱うことにより、広範囲のF/B比において、プロセッサ側からメモリシステム側にボトルネックの移行が起こる。

3.3 コンセプト

上記の観測結果に基づき、データサイズを考慮した電力シフティング手法を提案する。このコンセプトは図4に示されている。すなわち、プロセッサへの割り当て電力 P_{cpu} を下げることによって、プロセッサのスループット律速の水平線を下げつつ(多くのF/B比では性能低下が起きない)、それによって生じた余剰電力をメモリシステム側 P_{mem}, P_{mem2} に割り当てることで、大きいスケールのデータを利用することによって生じた性能低下をカバーするとい

うものである。ただし、条件 $P_{cpu} + P_{mem} + P_{mem2} \leq P_{total}$ は満たしつつ、これらを変化させるものとする。この様にして電力割り当てを工夫し、ルーフライン曲線をデータサイズに応じて調整し直すことで、ノードに与えられた電力予算 P_{total} を最大限に有効活用しようとするのが本提案である。

4. 評価

前述のデータサイズを考慮した電力シフティングの有効性を確認するため、さらなる予備評価を実施した。以下では、4.1章にて行った実験の内容を示し、4.2章にて実験結果を示す。

4.1 評価手法

表2に示すSKLベースのシステムにて図3のプログラムを動作させ、さらにCPUやメモリに対して電力制御を行った場合の性能を評価した。ただし、表1に示すKNLベースのシステムは、メモリの電力制御をサポートしていないため、本評価では利用していない(KNL側では電力制御機能をサポートしているものの[14]、BIOS側でロックがかかっているためである)。この電力制御のインターフェイスとしてRAPL(Running Average Power Limit)[15]を利用し、各コンポーネントに電力キャップを与えることで評価を行う。ただし、NVRAMの電力制御はここでは考え

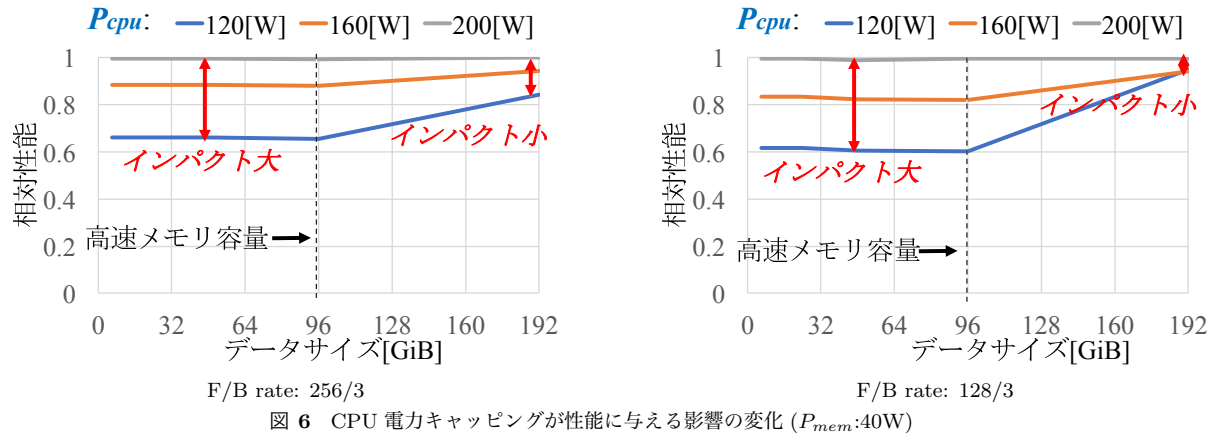


図 6 CPU 電力キャッピングが性能に与える影響の変化 ($P_{mem}:40W$)

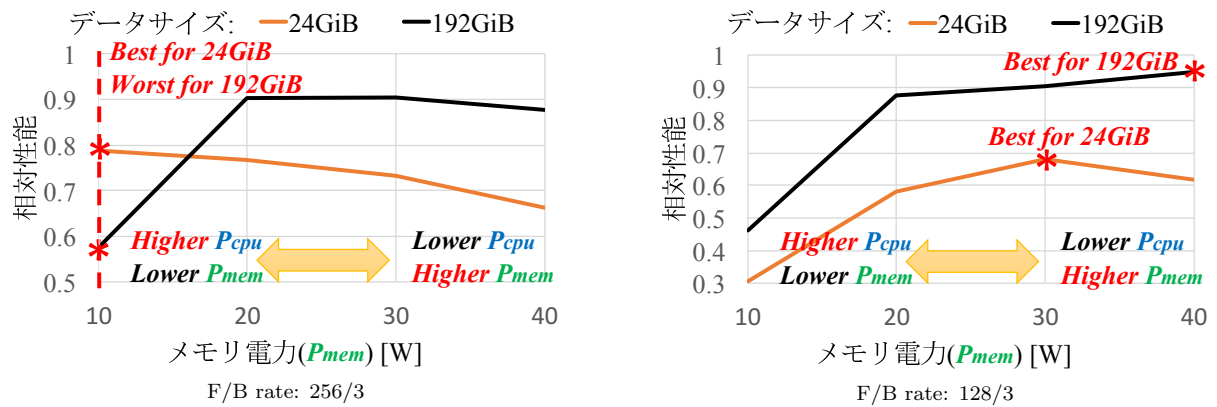


図 7 電力シフティングの結果 ($P_{cpu}+P_{mem}=P_{total}:160W$)

ないものとする (すなわち、 $P_{mem2} = 0$ とみなして評価を行う)。NVRAM の電力制御まで含めた評価については今後の研究課題である。

このような条件の下、以下の実験を行った。

- P_{mem} は固定値とし、 P_{cpu} をスケールさせる。 P_{total} の制約は設けず、 ∞ に設定する。
- P_{total} に固定値を与え、 P_{cpu}, P_{mem} を動かして、最適な $\{P_{cpu}, P_{mem}\}$ の組み合わせを探索する。

前者は、CPU の電力キャッピングの性能への影響が、データサイズを変化させることでどの様に変化するかを確かめることを目的とし、後者は、 $\{P_{cpu}, P_{mem}\}$ の最適組み合わせがデータサイズを変化させることでどの様に変化するかを確かめることを目的とする。

4.2 評価結果

図 6 に CPU 電力キャッピングが性能に与える影響に関する評価結果を示す。横軸はデータサイズであり、縦軸は相対性能 (各々のデータサイズにおける、電力制約を与えない場合の性能で正規化) である。図に示す通り、データサイズが大きくなればなるほど、CPU 電力キャップが性能に与える影響は小さくなっている。これは、図 4 に示す通り、データサイズが大きい場合には性能はプロセッサのスループットよりもメモリシステムのバンド幅律速となる

ことが多いためである。

図 7 に CPU と DRAM メモリの間での電力シフティングを行った場合の性能を示す。ただし、横軸は DRAM メモリに与えた電力キャップ値 P_{mem} を示しており、縦軸は相対性能 (各々のデータサイズにおける、電力制約を与えない場合の性能で正規化) を表している。ここで、CPU の電力キャップ値は $P_{cpu} = P_{total} - P_{mem}$ の様に与えている。図に示す通り、最適な $\{P_{cpu}, P_{mem}\}$ の最適組み合わせはデータサイズに応じて大きく変化している。特に F/B 比が 256/3 の場合に着目すれば、24GiB の際に最適な組み合わせだったものが、196GiB の際には最悪の組み合わせに変化している。このような結果から、データサイズを考慮した電力制御が有効であることが確認できた。

5. まとめと今後の方針

本稿では、ハイブリッドメモリを搭載するノードにおける、データサイズに着目した電力配分手法を提案し、予備評価によってその有効性を確認した。例えば、データサイズが変化することで、最適な電力配分であったものが、最悪の電力配分に変化することがあることを確認した。

今後の方針としては、まず、HPC ベンチマーク、実アプリケーション、さらにはデータセンタ向けベンチマーク等を利用して、本提案の有効性をより現実的な状況で確認

することが挙げられる。さらには、最適電力割り当てを行うためのアルゴリズムやフレームワークの作成を行うことも必要である。また、本稿では、NVRAM の電力マネジメントは考慮していないが、これを含めて電力割り当てを最適化することも重要である。さらには、Intel Knights Landing 等のシステムの上でも有効性を評価することも検討している。

謝辞 本研究の一部は、MEXT 科研費 JP16H06677 及び“ポスト京のプロセッサアーキテクチャ、電力制御技術、システムソフトウェア及び数値計算ライブラリに関する研究”の助成を受けたものである。

参考文献

- [1] Akinaga, H. and Shima, H.: Resistive Random Access Memory (ReRAM) Based on Metal Oxides, *Proceedings of the IEEE*, Vol. 98, No. 12, pp. 2237–2251 (2010).
- [2] Arima, E., Noguchi, H., Nakada, T., Miwa, S., Takeda, S., Fujita, S. and Nakamura, H.: Immediate Sleep: Reducing Energy Impact of Peripheral Circuits in STT-MRAM Caches, *ICCD*, pp. 149–156 (2015).
- [3] Bailey, P. E., Lowenthal, D. K., Ravi, V., Rountree, B., Schulz, M. and d. Supinski, B. R.: Adaptive Configuration Selection for Power-Constrained Heterogeneous Systems, *ICPP*, pp. 371–380 (2014).
- [4] Bourzac, K.: Has Intel Created a Universal Memory Technology? [News], *IEEE Spectrum*, Vol. 54, No. 5, pp. 9–10 (2017).
- [5] Cao, T., He, Y. and Kondo, M.: Demand-Aware Power Management for Power-Constrained HPC Systems, *CC-Grid*, pp. 21–31 (2016).
- [6] Consortium, H. M. C.: Hybrid Memory Cube Specification 2.1, *Last Revision Jan* (2015).
- [7] Deng, Q., Meisner, D., Bhattacharjee, A., Wensch, T. F. and Bianchini, R.: CoScale: Coordinating CPU and Memory System DVFS in Server Systems, *MICRO*, pp. 143–154 (2012).
- [8] Dhiman, G., Ayoub, R. and Rosing, T.: PDRAM: A Hybrid PRAM and DRAM Main Memory System, *DAC*, pp. 664–669 (2009).
- [9] Ellsworth, D. A., Malony, A. D., Rountree, B. and Schulz, M.: Dynamic Power Sharing for Higher Job Throughput, *SC*, pp. 80:1–80:11 (2015).
- [10] Fan, X., Weber, W.-D. and Barroso, L. A.: Power Provisioning for a Warehouse-sized Computer, *ISCA*, pp. 13–23 (2007).
- [11] Felter, W., Rajamani, K., Keller, T. and Rusu, C.: A Performance-conserving Approach for Reducing Peak Power Consumption in Server Systems, *ICS*, pp. 293–302 (2005).
- [12] Ge, R., Zou, P. and Feng, X.: Application-Aware Power Coordination on Power Bounded NUMA Multicore Systems, *ICPP*, pp. 591–600 (2017).
- [13] Hanson, H., Felter, W., Huang, W., Lefurgy, C., Rajamani, K., Rawson, F. and Silva, G.: Processor-Memory Power Shifting for Multi-Core Systems, *International Business Machines*, pp. 1–7 (2012).
- [14] Intel: Intel® 64 and ia-32 Architectures Software Developer’s Manual, Model-Specific Registers (2017).
- [15] Intel: Intel® 64 and ia-32 Architectures Software Developer’s Manual, System Programming Guide (2017).
- [16] Intel: Intel® Memory Drive Technology, Set Up and Configuration Guide (2017).
- [17] Jeffers, J., Reinders, J. and Sodani, A.: *Intel Xeon Phi Processor High Performance Programming: Knights Landing Edition*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2016).
- [18] Kaxiras, S., Hu, Z. and Martonosi, M.: Cache Decay: Exploiting Generational Behavior to Reduce Cache Leakage Power, *ISCA*, pp. 240–251 (2001).
- [19] Kim, J. and Kim, Y.: HBM: Memory Solution for Bandwidth-Hungry Processors, *Hot Chips 26 Symposium (HCS)*, pp. 1–24 (2014).
- [20] Komoda, T., Hayashi, S., Nakada, T., Miwa, S. and Nakamura, H.: Power Capping of CPU-GPU Heterogeneous Systems through Coordinating DVFS and Task Mapping, *ICCD*, pp. 349–356 (2013).
- [21] Lee, B. C., Ipek, E., Mutlu, O. and Burger, D.: Architecting Phase Change Memory As a Scalable Dram Alternative, *ISCA*, pp. 2–13 (2009).
- [22] Lefurgy, C., Wang, X. and Ware, M.: Power capping: a prelude to power shifting, *Cluster Computing*, Vol. 11, No. 2, pp. 183–195 (2008).
- [23] Li, J., Huang, W., Lefurgy, C., Zhang, L., Denzel, W. E., Treumann, R. R. and Wang, K.: Power Shifting in Thrifty Interconnection Network, *HPCA*, pp. 156–167 (2011).
- [24] Miwa, S. and Nakamura, H.: Profile-based Power Shifting in Interconnection Networks with on/off Links, *SC*, pp. 37:1–37:11 (2015).
- [25] Noguchi, H., Ikegami, K., Takaya, S., Arima, E., Kushida, K., Kawasumi, A., Hara, H., Abe, K., Shimomura, N., Ito, J., Fujita, S., Nakada, T. and Nakamura, H.: 7.2 4Mb STT-MRAM-Based Cache with Memory-Access-Aware Power Optimization and Write-Verify-Write / Read-Modify-write scheme, *ISSCC*, pp. 132–133 (2016).
- [26] Park, H., Yoo, S. and Lee, S.: Power Management of Hybrid DRAM/PRAM-Based Main Memory, *DAC*, pp. 59–64 (2011).
- [27] Patki, T., Lowenthal, D. K., Rountree, B., Schulz, M. and de Supinski, B. R.: Exploring Hardware Overprovisioning in Power-constrained, High Performance Computing, *ICS*, pp. 173–182 (2013).
- [28] Paul, I., Huang, W., Arora, M. and Yalamanchili, S.: Harmonia: Balancing Compute and Memory Power in High-performance GPUs, *ISCA*, pp. 54–65 (2015).
- [29] Ramos, L. E., Gorbato, E. and Bianchini, R.: Page Placement in Hybrid Memory Systems, *ICS*, pp. 85–95 (2011).
- [30] Sakamoto, R., Cao, T., Kondo, M., Inoue, K., Ueda, M., Patki, T., Ellsworth, D., Rountree, B. and Schulz, M.: Production Hardware Overprovisioning: Real-World Performance Optimization Using an Extensible Power-Aware Resource Management Framework, *IPDPS*, pp. 957–966 (2017).
- [31] Sarood, O., Langer, A., Kal, L., Rountree, B. and de Supinski, B.: Optimizing Power Allocation to CPU and Memory Subsystems in Overprovisioned HPC Systems, *CLUSTER*, pp. 1–8 (2013).
- [32] Sarood, O., Langer, A., Gupta, A. and Kale, L.: Maximizing Throughput of Overprovisioned HPC Data Centers Under a Strict Power Budget, *SC*, pp. 807–818 (2014).
- [33] Savoie, L., Lowenthal, D. K., d. Supinski, B. R., Islam,

- T., Mohror, K., Rountree, B. and Schulz, M.: I/O Aware Power Shifting, *IPDPS*, pp. 740–749 (2016).
- [34] Standard, J.: High Bandwidth Memory (HBM) DRAM, *JESD235* (2013).
- [35] Vijayaraghavan, T., Eckert, Y., Loh, G. H., Schulte, M. J., Ignatowski, M., Beckmann, B. M., Brantley, W. C., Greathouse, J. L., Huang, W., Karunanithi, A., Kayiran, O., Meswani, M., Paul, I., Poremba, M., Raasch, S., Reinhardt, S. K., Sadowski, G. and Sridharan, V.: Design and Analysis of an APU for Exascale Computing, *HPCA*, pp. 85–96 (2017).
- [36] Williams, S., Waterman, A. and Patterson, D.: Roofline: An Insightful Visual Performance Model for Multicore Architectures, *Communications of the ACM*, Vol. 52, No. 4, pp. 65–76 (2009).
- [37] Zhang, J., Liao, X., Jin, H., Liu, D., Lin, L. and Zhao, K.: An Optimal Page-Level Power Management Strategy in PCM—DRAM Hybrid Memory, *Int. J. Parallel Program.*, Vol. 45, No. 1, pp. 4–16 (2017).